Multi-instantons and supersymmetric SU(N) gauge theories

Pomeroy, Neil B.

Multi-Instantons and Supersymmetric $SU(N)$ Gauge Theories

A thesis submitted for the degree of
Doctor of Philosophy
by

Neil B. Pomeroy

University of Durham
Department of Physics
Institute for Particle Physics Phenomenology

2002

Acknowledgements

I wish to thank Dr. Valya Khoze for his encouragement and guidance, and for his inspiration and patience, as my supervisor. The numerous key theoretical and mathematical insights which he has provided for me have been vital to my progress during this work.

My deepest gratitude to Mum, Dad, and my brother Adam, without whose continuous support and sound advice I would have been unable to make it through the brightest and darkest times I have experienced in the past three years. Thank you for always listening to me, whatever I had to say.

To my office friends, Lee Garland and Kate Adamson: thank you. I could not imagine more pleasant and amiable colleagues with whom to share an office and research in physics with for three years.

My thanks to: Dr. Chris Maxwell, Prof. W. James Stirling, and Prof. Mike Pennington, without whom the opportunity to become a research student would not have come about; Dr. E. W. Nigel Glover, for invaluable advice; Dr. Clifford Johnson, Prof. David Fairlie, Dr. Gabriele Travaglini, and Dr. Peter Bowcock, for discussions and advice; Dr. Mike Leech for solving many of my computer problems; and finally, Linda Wilkinson, the Physics Office staff, and Mike Lee for excellent stationary support.

I thank the U.K. Particle Physics and Astronomy Research Council for funding my research studentship.
My thanks to those who have gone before: Dr. N. Michael Davies, Athanosios Koukoutsakis, Dr. Charalampos Anastasiou, Dr. Maria-Elena Tejada-Yeomans, Dr. Mark Hale, Dr. David Hickin and Dr. Vinay Kotecha. I must also mention Dr. Matthew Slater, whose brief but influential discussions with me in 1999 were more far reaching than I realized at the time.

'Virtutis Fortuna Comes'

'Etiamsi omnes, ego non'

Declaration

I declare that no material presented in this thesis has previously been submitted for a degree at this or any other university.

The research described in this thesis was performed under the guidance of Dr. V. V. Khoze, and has been published as follows:

1. N. B. Pomeroy, *Matching One-Instanton Predictions and Exact Results in $\mathcal{N} = 2$ Supersymmetric $SU(N)$ Theories*, Phys. Lett. B **501** (2001) 511, [arXiv: hep-th/0103181],

2. N. B. Pomeroy, *The $U(N)$ ADHM Two-Instanton*, Phys. Lett. B **547** (2002) 85, [arXiv: hep-th/0203184].

The contributions of the author are described in Subsections 2.3.2 and 2.3.3 of Chapter 2, and Subsection 6.5.1 of Chapter 6.

Abstract

In this thesis the proposed exact results for low energy effective $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory coupled to N_f fundamental matter multiplets in four dimensions are considered. The proposed exact results are based upon the work of Seiberg and Witten for low energy effective four dimensional $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory coupled to N_f fundamental matter multiplets. The testing and matching of the proposed exact results via supersymmetric instanton calculus are the motivation for two studies. Firstly, we study the ADHM construction of instantons for gauge groups $U(N)$ and $SU(2)$ and for topological charge two and three. The ADHM constraints which implicitly specify instanton gauge field configurations are solved for the explicit exact general form of instantons with topological charge two and gauge group $U(N)$. This is the first explicit and general multi-instanton configuration for the unitary gauge groups. The $U(N)$ ADHM two-instanton configuration may be used in further tests and matching of the proposed exact results in low energy effective $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theories by comparison with direct instanton calculations.

Secondly, a one-instanton level test is performed for the reparameterization scheme proposed by Argyres and Pelland matching the conjectured exact low energy results and instanton predictions for the instanton contributions to the prepotential of low energy effective $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with $N_f = 2N$ massless fundamental matter multiplets. The constants within the reparameterization scheme which ensure agreement between the exact results and the instanton predictions for general $N > 1$ are derived for the entire quantum moduli space. This constitutes a non-trivial test of the proposed reparameterization scheme, which eliminates the discrepancies arising when the two sets of results are compared.

Acknowledgements

I wish to thank Dr. Valya Khoze for his encouragement and guidance, and for his inspiration and patience, as my supervisor. The numerous key theoretical and mathematical insights which he has provided for me have been vital to my progress during this work.

My deepest gratitude to Mum, Dad, and my brother Adam, without whose continuous support and sound advice I would have been unable to make it through the brightest and darkest times I have experienced in the past three years. Thank you for always listening to me, whatever I had to say.

To my office friends, Lee Garland and Kate Adamson: thank you. I could not imagine more pleasant and amiable colleagues with whom to share an office and research in physics with for three years.

My thanks to: Dr. Chris Maxwell, Prof. W. James Stirling, and Prof. Mike Pennington, without whom the opportunity to become a research student would not have come about; Dr. E. W. Nigel Glover, for invaluable advice; Dr. Clifford Johnson, Prof. David Fairlie, Dr. Gabriele Travaglini, and Dr. Peter Bowcock, for discussions and advice; Dr. Mike Leech for solving many of my computer problems; and finally, Linda Wilkinson, the Physics Office staff, and Mike Lee for excellent stationary support.

I thank the U.K. Particle Physics and Astronomy Research Council for funding my research studentship.
My thanks to those who have gone before: Dr. N. Michael Davies, Athanosios Koukoutsakis, Dr. Charalampos Anastasiou, Dr. Maria-Elena Tejada-Yeomans, Dr. Mark Hale, Dr. David Hickin and Dr. Vinay Kotecha. I must also mention Dr. Matthew Slater, whose brief but influential discussions with me in 1999 were more far reaching than I realized at the time.

'Virtutis Fortuna Comes'

'Etiamsi omnes, ego non'

Declaration

I declare that no material presented in this thesis has previously been submitted for a degree at this or any other university.

The research described in this thesis was performed under the guidance of Dr. V. V. Khoze, and has been published as follows:

1. N. B. Pomeroy, *Matching One-Instanton Predictions and Exact Results in $\mathcal{N} = 2$ Supersymmetric $SU(N)$ Theories*, Phys. Lett. B **501** (2001) 511, [arXiv: hep-th/0103181],

2. N. B. Pomeroy, *The $U(N)$ ADHM Two-Instanton*, Phys. Lett. B **547** (2002) 85, [arXiv: hep-th/0203184].

The contributions of the author are described in Subsections 2.3.2 and 2.3.3 of Chapter 2, and Subsection 6.5.1 of Chapter 6.

Contents

Chapter 1

Introduction

Quantum field theory is the central tool of theoretical high energy physics. Quantum gauge field theories in four dimensional spacetime provide the current models for the fundamental interactions between elementary particles. The most important gauge field theories are Yang–Mills gauge field theories, in which the gauge group is non-Abelian. The Standard Model of particle physics is based upon gauge theories, with quantum Yang–Mills gauge field theories assuming a vital rôle.

Perturbation theory is a method by which the experimentally verifiable predictions can be derived from quantum gauge field theories. In perturbation theory, one expands an analytic function of the gauge coupling when the gauge coupling assumes small values; the theory is referred to as being in the weakly coupled regime for these values. Weak coupling enables one to make a valid expansion of functions dependent on the coupling in quantum field theory. At strong coupling, perturbations in the gauge coupling cannot be made, and no information regarding the strong coupling regime may be extracted using this method.

Phenomenological applications of quantum gauge field theory almost exclusively use perturbation theory to obtain quantitative predictions from these models. However, quantum field theories also contain functions which are non-analytic in the gauge coupling constant. Perturbation theory cannot be applied to such functions, with the result that perturbative techniques do not yield any results about the physical content of these functions. Hence functions which are non-analytic in the gauge coupling must be treated non-perturbatively. Non-perturbative methods in quantum gauge field theory present many difficulties. Analytical non-perturbative methods use the semi-classical approximation of the path integral formulation of quantum field theory. In the semi-classical approximation, quantum fluctuations about exact non-trivial classical solution of the equations of motion of the field theory are used to evaluate the non-perturbative contributions which the solution makes to the path integral. A functional expansion in the fields about these classical solutions can be used to derive non-perturbative information, in the form of functions which are non-analytic in the gauge coupling, about the theory under consideration. Such an expansion is only valid at weak coupling. Non-perturbative effects vanishing at strong coupling, and so weak coupling must be used. It is convenient to change from Minkowski spacetime to Euclidean spacetime when considering the non-perturbative effects in a quantum field theory. In Euclidean Yang–Mills gauge theories, at weak coupling, the dominant non-trivial classical solutions are known as instantons and anti-instantons. Instantons and anti-instantons are exact classical minima of the four dimensional Euclidean Yang–Mills action, and exist as exact solutions to the (anti-)self-dual Yang–Mills field equations. All solutions to the (anti-)self-dual Yang–Mills field equations can be classified using a particular mathematical method. Instantons (that is, both instantons and anti-instantons) can be interpreted as fluctuations of the gauge field vacuum and as tunnelling processes which connect inequivalent vacuum states. There are an infinite number of instanton solutions which will contribute non-perturbatively to the quantum Yang–Mills gauge field theory, known as multi-instantons. These classical gauge field configurations cannot readily be obtained in explicit and general form. In this thesis we investigate multi-instanton solutions for the gauge group $SU(N)$.

Instantons contribute to phenomenologically valuable quantum Yang–Mills gauge field theories, such as those in the Standard Model, in a complicated way. All combinations of instantons and anti-instantons contribute, and are thought to be involved in the lifting of any vacuum state degeneracy in these theories. Non-perturbative contributions from instantons are negligible compared to those from perturbation theory, and approximate models must be used to evaluate their effects even on microscopic scales. In supersymmetric gauge theories, instanton effects can often be calculated exactly. Supersymmetry is a theoretical symmetry of field theories which has many appealing properties. Not only does supersymmetry appear to provide the solution to problems in the phenomenology of gauge field theories, such as the hierarchy problem, it also provides mathematically elegant theoretical models of elementary particles. Furthermore, supersymmetric gauge theories also exhibit important physical features in common with less symmetric gauge field theories, such as confinement and chiral symmetry breaking. Non-perturbative effects, such as those arising from instantons, can be exactly calculated in a number of supersymmetric quantum field theories due to the constraints supersymmetry places on quantum corrections. Instanton contributions are prominent in many supersymmetric gauge theories, which in general possess an infinite number of degenerate vacuum states, which instantons cannot rectify.

Many exact non-perturbative results have been obtained in quantum supersymmetric gauge theories using semi-classical methods. Less than ten years ago, Seiberg and Witten were able to exactly determine the low energy Wilsonian effective action of four dimensional $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge field theory. Conventional methods used previously did not yield this effective action. The solution for the low energy Wilsonian effective action is valid at both weak and strong coupling, and reveals the strongly coupled gauge field dynamics of the theory. Known perturbative and non-perturbative methods cannot be used to analyse the strongly coupled regime of quantum gauge field theories, and so the Seiberg–Witten solution presents major progress towards understanding the strong coupling behaviour of quantum gauge field theories. The Seiberg–Witten is also the first known *exact* solution of the low energy dynamics of a four dimensional quantum field theory.

The techniques used by Seiberg and Witten are not those of conventional field theory. Rather, they are a synthesis of physical intuition and various previously known results specific to $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge field theory. The low energy Wilsonian effective action of $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge field theory is completely specified by a holomorphic function known as the prepotential. Seiberg and Witten propose, via an elaborate sequence of arguments, an exact prepotential for the theory which meets all of the necessary criteria for this function. The prepotential is known to receive an infinite series of quantum non-perturbative corrections from instantons. The Seiberg–Witten solution proposes an exact evaluation of this series. Motivated by these developments, direct calculation of this series using semi-classical non-perturbative methods from first principles were performed. The presence of supersymmetry provides many simplifications in the analytic calculation of the instanton contributions to the prepotential. Comparisons made between the proposed exact non-perturbative contributions with semi-classical calculations made using known multi-instanton gauge field configurations are in agreement with some exceptions. The most serious discrepancies occur for $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theories coupled to $2N$ fundamental matter multiplets. These exceptions are significant, since the techniques used by Seiberg and Witten have found widespread application in other areas of theoretical high energy physics, notably string theory, and also mathematical physics and pure mathematics. But the exact results proposed by Seiberg and Witten, and their subsequent generalization to other $\mathcal{N} = 2$ supersymmetric gauge field theories, are highly non-trivial solutions which contain information other than the exact $\mathcal{N} = 2$ prepotential. However, in order to agree with the results of conventional non-perturbative calculations, the proposed exact results must be matched to the relevant instanton predictions. A recently proposed matching scheme purports to resolve all of the discrepancies between the two sets of results, through a non-perturbative reparameterization. This scheme generalizes other work for special cases of discrepancy. In this thesis we investigate the explicit matching within this scheme between the proposed exact results and instanton calculations for $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory coupled to $2N$ fundamental matter multiplets.

In Chapter 2 we review instantons in Yang–Mills field gauge theories. We begin with a general description of how instantons arise in Yang–Mills gauge field theories. Instantons are exact solutions of the classical (anti-)self-dual Yang–Mills field equations which give a minimum finite action of Yang–Mills gauge field theory in four dimensional Euclidean spacetime. It is shown that instanton gauge field configurations can be classified by an integer referred to as the topological charge or instanton number. The instanton number is used to prefix the particular instanton configuration under consideration. Multi-instanton configurations are instanton solutions with topological charge greater than one. We then describe the first known instanton configuration, discovered by Belavin, Polyakov, Schvarts and Tyupkin (BPST), and named the BPST instanton. This is the most general one-instanton solution of pure $SU(2)$ Yang–Mills gauge field theory. We then describe the concept of the instanton moduli space, which is important for characterizing instanton solutions. This is followed by a detailed description of a general method for constructing instanton configurations formulated by Atiyah, Drinfeld, Hitchin and Manin (ADHM), known as the ADHM construction of instantons. This method implicitly defines all multi-instanton solutions for Yang–Mills gauge theories with arbitrary classical gauge groups. We describe the ADHM construction for gauge groups $U(N)$ and $Sp(N)$. In order to extract explicit multi-instanton solutions from the ADHM construction, a set of non-trivial constraints, known as the ADHM constraints, are to be solved. The ADHM constraints are solved to give the general $U(N)$ two-instanton solution. The ADHM constraints for the $U(N)$ three-instanton and the $Sp(N)$ three-instanton are also described.

Supersymmetric gauge theories are reviewed in Chapter 3. The concept of global supersymmetry is described. The degree of supersymmetry in a supersymmetric gauge theory is specified by an integer $\mathcal{N}$, which indicates the number of supersymmetry generators present. The notion of supersymmetry constraints is briefly outlined. This is followed by a description of the $\mathcal{N} = 1$ superfield formalism, which can be used to efficiently construct supersymmetric gauge theories. We then describe $\mathcal{N} = 1$ and $\mathcal{N}$-extended supersymmetric Yang–Mills gauge theories, up to the maximally extended case of $\mathcal{N} = 4$ supersymmetric gauge theories.

In Chapter 4 we review exact results in $\mathcal{N} = 1$ and $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theories. Methods which make use of instantons have previously been used to obtain exact results in supersymmetric field theories. We then describe the concept of duality in $\mathcal{N} = 1$ and $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theories. Duality between the electric and magnetic degrees of freedom and between the weak and strong coupling regimes is a property conjectured to be present in $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theories. In particular, we describe the generalization of electric-magnetic duality known as S-duality in $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge field theory. Prior to this we briefly review magnetic monopoles in Yang–Mills gauge theories, which are central to the realization of electric-magnetic duality in field theories. We then describe a special form of duality in $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theories, known as Seiberg duality. Seiberg duality is not equivalent to electric-magnetic duality, and is restricted to $\mathcal{N} = 1$ supersymmetric gauge theories, but it does involve correspondences between phenomenon at weak and strong coupling, and electric and magnetic degrees of freedom. Before this, we review the concepts of the moduli space of vacua and phases of $\mathcal{N} = 1$ supersymmetric gauge theories.

Many of the concepts described in Chapter 4 appear again in the context of $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theories. These include electric-magnetic duality and the concept of a moduli space of vacua. In Chapter 5, we first describe exact results for $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theories analogous to those obtained for $\mathcal{N} = 1$ and $\mathcal{N} = 4$ supersymmetric gauge theories. We then present a detailed review of Seiberg–Witten theory in Chapter 5. Seiberg–Witten theory proposes that the low energy dynamics of quantum $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory and quantum $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD in the Coulomb phase can be determined from reasoning involving electric-magnetic duality, $\mathcal{N} = 2$ supersymmetry and the moduli space of vacua in these theories. The low energy dynamics of these theories are specified by the prepotential, a holomorphic function of the superfields. Seiberg and Witten propose that the prepotential can be reconstructed exactly from knowledge of the moduli space of vacua via complex analysis. The exact form of the low energy effective action of the theory, valid at both strong and weak coupling, can then be proposed. We follow this with a description of the generalizations of the methods of Seiberg–Witten theory to apply to $\mathcal{N} = 2$ supersymmetric gauge theories with other gauge groups and matter contents. In particular we focus upon $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD, which is the theory of primary interest in this thesis.

In Chapter 6 we describe the application of instanton calculus to test and match the proposed exact results in $\mathcal{N} = 2$ supersymmetric QCD with gauge group $SU(2)$ and $SU(N)$. We first review instanton calculus in Yang–Mills gauge theory and supersymmetric Yang–Mills gauge theory. To calculate the quantum effects of instantons, the semi-classical approximation is employed. The collective coordinate method, which makes use of the ADHM construction of instantons, is also used in instanton calculus to enable calculations to be performed. In the Coulomb phase of $\mathcal{N} = 2$ supersymmetric gauge theories, the gauge symmetry is broken by the non-zero vanishing expectation values of scalar fields. In this way, the fermion fields present in supersymmetric gauge theories affect the bosonic pure Yang–Mills gauge theory, of which instantons are classical solutions. Due to this, instantons are no longer exact solutions of the classical field equations of supersymmetric gauge theories. Despite this, instanton effects can still be calculated using the constrained instanton formalism. We then state the results of the comparison of instanton calculations with the results of Seiberg–Witten theory for the instanton contributions to various quantities in $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD. These consist of a one-instanton and a two-instanton test of the instanton contributions to the prepotential, and a special all-orders instanton test of a renormalization group relation. An extension of the one-instanton test of Seiberg–Witten theory to the case of the prepotential of $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD is then described. We then describe the discrepancies which have been found between the proposed exact results and instanton predictions as a result of instanton tests of Seiberg–Witten theory and its generalizations. The most serious discrepancy occurs in $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD with $N_f = 2N$ fundamental matter multiplets, of which the Seiberg–Witten result for $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD is a special case. Following earlier suggestions, a unified reparameterization scheme, known as the Argyres–Pelland matching scheme, has been proposed to resolve these discrepancies. Through a non-perturbative reparameterization, quantities such as the gauge coupling used in the proposed exact results can be matched to the same quantities derived from instanton calculations. Working within this scheme, we consider the matching between the proposed exact result for the one-instanton contribution to the prepotential of $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD with $N_f = 2N$ massless fundamental matter multiplets and the one-instanton calculation for the same quantity. Precise agreement between these two sets of results is obtained by matching the one-instanton predictions from both sources for all $N > 1$. We present the formulae for the constants in the Argyres–Pelland matching scheme in this case, which are valid for all $N > 1$.

Upon first reading, the reader may omit Chapter 4, and in particular Sections 4.4 and

4.5 of Chapter 4, without any loss of essential material.

Our conclusion is given in Chapter 7, which is followed by two appendices and a bibliography of works cited or consulted.

Chapter 2

Yang–Mills Instantons

2.1 Introduction

Instantons have proved to be an important phenomenon in Yang–Mills gauge field theory ever since their discovery [1]. They have been the subject of a wealth of mathematical and theoretical literature over the last three decades. At weak coupling, instantons constitute the dominant non-perturbative effect in Yang–Mills gauge theories. Yang–Mills gauge theories are the theories upon which much of the phenomenologically successful models of particle physics are founded. Furthermore, instantons assume an unusual status in theoretical high energy physics: their existence and nature was deduced purely from the properties of classical Yang–Mills gauge theory, and they are considered physical processes, but experimentally they have never, and likely will never, be directly observed. Thirty years ago, it was hoped that confinement in Yang–Mills gauge theory could be attributed to instantons, but the puzzle of confinement has not allowed itself to be resolved so straightforwardly [15, 17]. However, instantons are still studied in the context of phenomenological theories such as QCD [70, 71, 72, 73] and observable physical effects can be attributed to them [74].

More recently, instantons have been revived as their effects can be exactly calculated in supersymmetric gauge theories. This follows the advances made in determining the exact low energy effective descriptions of such theories, which we shall return to later in this thesis. They have also been used in as a tool in pure mathematics for the description and study of the differential topology of four-manifolds. Recently instantons have also been generalized to noncommutative spacetimes and higher dimensions.

Instantons are essentially non-trivial exact solutions of a special form of the classical Yang–Mills field equations which locally minimize the Yang–Mills action in four dimensional Euclidean spacetime. Yang–Mills gauge fields with this property necessarily satisfy a set of field equations known as the self-dual Yang–Mills equations. Accordingly, instantons are self-dual gauge fields, which are topological in nature. They represent non-trivial fluctuations in the gauge field vacuum in four dimensional spacetime, and are localised in both space and time, making their existence transient. Hence their description as processes rather than particles is more appropriate. As we shall describe below, instantons can be interpreted as processes connecting vacua in a Yang–Mills gauge field theory via quantum mechanical tunnelling.

In Section 2.2, we describe instantons in pure Yang–Mills gauge field theory and their properties. We describe the first known instanton solution, found by Belavin, Polyakov, Schvarts and Tyupkin (BPST) [1] for Yang–Mills gauge field theory in Subsection 2.2.1. The properties of the BPST instanton and instantons in general are also described in Subsection 2.2.1. In Subsection 2.2.2 we describe the collective co-ordinates and the moduli space of instantons.

Using instanton configurations in quantum gauge field theory requires the instanton calculus. The instanton calculus shall be described in Chapter 6, where applications of instanton calculations are also described. Mathematical descriptions of all general instanton configurations with minimum finite Yang–Mills actions in Euclidean spacetime are usually necessary for instanton calculations. To construct these descriptions of instantons, one can use the construction of Atiyah, Drinfeld, Hitchin and Manin (ADHM) for instantons [8]. The ADHM construction is a technique for describing all self-dual gauge fields, with arbitrary classical gauge group and instanton charge. In Section 2.3 we describe the ADHM construction. In Subsection 2.3.1 we focus on the ADHM construction for instantons with gauge group $U(N)$ and describe the first exact general multi-instanton solution found for $U(N)$ Yang–Mills gauge theory. This is the exact general $U(N)$ two-instanton solution found by solving the appropriate ADHM constraints. In Subsection 2.3.3 we describe the ADHM construction for instantons with gauge group $Sp(N)$. The construction of $Sp(N)$ ADHM instantons can be approached as a special case of the $U(N)$ construction or through its own formalism. The ADHM construction of $Sp(N)$ instantons is more simple than that for $U(N)$ instantons. This fact, in conjunction with the isomorphism $Sp(1) \simeq SU(2)$, enables one to describe $SU(2)$ ADHM multi-instantons in a more efficient and economical manner than by using the $SU(N)$ formalism.

2.2 Instantons in Yang–Mills Gauge Theory

For definiteness and later reference, we describe instantons in four dimensional $SU(N)$ Yang–Mills gauge field theory in the first part of this section. Later, in Subsection 2.2.1 we will specialise to $SU(2)$ Yang–Mills gauge fields. In this section we make use of the reviews [49, 50, 63, 64, 75, 76, 90, 224, 280, 282, 283]. We also refer to the original papers concerning instantons in Yang–Mills gauge field theory [1, 2, 15, 16, 17, 18, 19].

We denote the gauge field as the function $v_m(x)$, where $m = 0, \ldots, 3$ is the Lorentz spacetime index and $x = x^m$ is the spacetime co-ordinate. The gauge field v_m is a Yang–Mills gauge field if the gauge group under which it is invariant is a non-Abelian group. A Yang–Mills gauge field can be decomposed as a matrix product of the physical gauge field v_m^a and the generators of the gauge group T^a, so that:

$$v_m(x) \;=\; v_m^a(x) T^a. \tag{2.1}$$

As is conventional in instanton literature, the gauge field v_m is taken to be an anti-Hermitian field. The covariant derivative in the field theory then acts as $D_m f = \partial_m f + g[v_m, f]$, with no factor of $i = \sqrt{-1}$ multiplying v_m, for some function f. The gauge field strength for a non-Abelian or Yang–Mills gauge field is then given by:

$$v_{mn} \;=\; \partial_m v_n - \partial_n v_m + g[v_m, v_n], \tag{2.2}$$

where g is the gauge coupling constant.

Instantons are phenomena which occur in four dimensional Euclidean spacetime. Throughout this chapter we work in four dimensional Euclidean spacetime. In Appendix A we detail our conventions for Euclidean spacetime. The action of the $SU(N)$ Yang–Mills gauge theory in Euclidean spacetime has the form:

$$S[v_m] \;=\; -\frac{1}{2} \int d^4x \, \mathrm{tr}_N \left(v_{mn} v^{mn} \right) - i\vartheta k, \tag{2.3}$$

where ϑ is a real number referred to as the vacuum angle or ϑ-angle, tr_N denotes the trace over the $SU(N)$ gauge group index and k is an integer termed the topological charge. The term $i\vartheta k$ is referred to as the 'theta term' or ϑ-term. The topological charge is an integer number, $k \in \mathbb{Z}$, given by the formula:

$$k = -\frac{g^2}{16\pi^2} \int d^4x \, \mathrm{tr}_N \, v^*_{mn} v^{mn}, \tag{2.4}$$

where $^*v_{mn}$ is dual of the gauge field strength v_{mn}. The dual gauge field strength is defined by:

$$^*v_{mn} \equiv \tfrac{1}{2}\epsilon_{mnkl} v_{kl}. \tag{2.5}$$

Equation (2.4) has the form of a total derivative. This is the origin of the ϑ-term in the action $S[v_m]$ in Eq. (2.3). Since $i\vartheta k$ is a total derivative, it can be added to the rest of the action in $S[v_m]$ with no effect on the equations of motion of the theory. A factor involving the ϑ-term will also be present in the partition function of the theory.

We now describe the origin of the topological charge k given in Eq. (2.4), which is an important parameter of instanton configurations. If the action Eq. (2.3) is to be finite, a necessary condition on the gauge field strength is that it vanishes at infinitely long range:

$$\lim_{|x|\to\infty} v_{mn} = 0. \tag{2.6}$$

It follows that a necessary and sufficient condition on the gauge field v_m which ensures that v_{mn} satisfies Eq. (2.6) is:

$$\lim_{|x|\to\infty} v_m = \frac{i}{g} U \partial_m U^{-1}, \tag{2.7}$$

where $U(x) \in G$ is an element of the gauge group $G = SU(N)$. The condition Eq.(2.7) appears to imply that at large distances, the gauge field v_m must tend to a gauge transformation of the classical, trivial vacuum $v_m = 0$, or 'pure gauge', for the gauge field strength v_{mn} to satisfy Eq. (2.6). If this were so, then the group element U, which we take to be a matrix, is the same as the gauge group element Ω which appears in ordinary local gauge transformations, in which:

$$v_m \to v'_m = \Omega v_m \Omega^{-1} + \frac{i}{g} \Omega \partial_m \Omega^{-1}. \tag{2.8}$$

The condition Eq. (2.7) does not demand that the matrix U is equivalent to the matrix Ω. The matrix Ω represents a continuous mapping from the field space, which is

Euclidean spacetime $\mathbb{R}^4$, to the gauge group $G = SU(N)$. This mapping can always be continuously deformed to the trivial mapping from $\mathbb{R}^4$ to a single element of G. Unlike Ω, the matrix U represents a continuous mapping from the 3-sphere S^3 at infinity, denoted S^3_∞ to the gauge group G. This mapping cannot in general be continuously deformed to the trivial mapping between $\mathbb{R}^4$ and G. In its most simple form, the matrix U acts as a map between S^3_∞ and, for example, the gauge group $G = SU(2)$ is isomorphic to S^3, written as $SU(2) \simeq S^3$, where $\simeq$ denotes isomorphism. The set of all continuous mappings between two 3-spheres can be categorised into different equivalence classes, such that in each class the elements can be continuously deformed into each one another. These equivalence classes can be labelled by an integer number $\kappa \in \mathbb{Z}$, which indicates the number of times the 3-spheres 'wind' about each other in the mapping, giving a measure of the non-trivial topology of the map. (We denote this integer k as it shall be identified with the 'instanton number', which is an integer measuring the instanton action and the number of constituent one-instantons in a k-instanton configuration.) The equivalence classes so labelled form a group, known as a homotopy group. The integer κ is known as the 'winding number', and can be shown to be exactly equivalent to the topological charge k defined in Eq. (2.4). Thus we write $\kappa = k$ hereafter and denote k as the winding number. In terms of differential topology, when the spacetime dimension is even, k is related to the Pontryagin index; when the spacetime dimension is odd, $k > 2$ is related to the second Chern class [66]. When $k = 0$, U can be deformed to the trivial mapping between $\mathbb{R}^4$ and $G = SU(2)$, when there are no windings, and one has $U \simeq \Omega$. Otherwise, when $k \neq 0$, different regions, or sectors, of the configuration space of finite actions are specified. Instantons are the gauge field configurations which minimize the action $S[v_m]$ in each different sector labelled by the topological charge.

For general gauge groups, $G \neq SU(2)$, Bott's theorem [11] states that for continuous mappings between S^3 and an arbitrary simple Lie group G, the mapping can be continuously deformed into a mapping between S^3 and an $SU(2)$ subgroup of G. Then the configuration space of finite actions can be labelled by the winding number for U as a mapping between S^3_∞ and $SU(2) \subset G$ subgroups of G. Instanton configurations can then be distinguished by their topological charge k, or 'instanton charge' or 'instanton number' as it is known in theoretical physics literature. That the topological charge is identical to the instanton charge can be established by substituting an instanton configuration into Eq. (2.4).

We now describe gauge field configurations which minimize the Yang–Mills action $S[v_m]$ given by Eq. (2.3). To be physically valid, we require that the action is more than or equal to zero. The Euclidean action cannot be less than zero; it is bounded from below. One can write the real part of the action $S[v_m]$ in terms of the gauge field strength v_{mn} and its dual ${}^*v^{mn}$ as follows:

$$\begin{aligned} S[v_m] = & -\frac{1}{4}\int d^4x\, \mathrm{tr}_N\left[(v_{mn} \pm {}^*v^{mn})^2\right] \pm \frac{1}{2}\int d^4x\, \mathrm{tr}_N\left(v_{mn}{}^*v^{mn}\right) \geq 0, \\ \Rightarrow & -\frac{1}{2}\int d^4x\, \mathrm{tr}_N\left(v_{mn} \pm {}^*v^{mn}\right)^2 \geq 0, \end{aligned} \tag{2.9}$$

where we have made use of the property ${}^*v^{mn}{}^*v_{mn} = v^{mn}v_{mn}$ in Euclidean spacetime. Using the formula for the topological charge k in Eq. (2.4), the following lower bound on the real part of $S[v_m]$ can be derived from Eq. (2.9):

$$-\frac{1}{2}\int d^4x\, \mathrm{tr}_N\left(v_{mn}v^{mn}\right) \geq \frac{8\pi^2}{g^2}|k|. \tag{2.10}$$

This bound becomes an equality and the action is locally minimized when the gauge field strength is either self-dual or anti-self-dual; that is, when the first right hand side term in Eq. (2.9) vanishes and v_{mn} satisfies:

$$v_{mn} = \pm{}^*v_{mn} \equiv \pm\tfrac{1}{2}\epsilon_{mnkl}v_{kl}. \tag{2.11}$$

The equations contained in Eq. (2.11) are known as the (anti-)self-dual Yang–Mills field equations, and are a set of non-linear first order partial differential equations for the gauge field v_m. The gauge fields v_m whose gauge field strengths satisfy Eq. (2.11) are referred to as self-dual or anti-self-dual gauge fields. Whether a gauge field is self-dual or anti-self-dual determines the sign of the topological charge of the gauge field:

$$v_{mn} = {}^*v_{mn}, \quad k > 0, \tag{2.12}$$

$$v_{mn} = -{}^*v_{mn}, \quad k < 0, \tag{2.13}$$

where for self-dual gauge fields, $k > 0$, and for anti-self-dual gauge fields $k < 0$. Both self-dual and anti-self-dual gauge fields locally minimize the four dimensional Euclidean Yang–Mills action. Self-dual gauge field configurations are referred to as instantons; anti-self-dual gauge field configurations are referred to as anti-instantons. This division follows from the positivity of the action $S[v_m]$: self-dual gauge fields with negative topological charge k would lead to a negative lower bound on the action, which is unphysical. Therefore only self-dual gauge fields with positive k can satisfy the self-dual Yang–Mills field equations. In this thesis we shall work exclusively with instantons; all references, except where stated otherwise, to 'instantons', will mean self-dual gauge fields with topological charge $k > 0$.

Gauge fields which obey Eq. (2.11) also automatically satisfy the Euler–Lagrange equations of the action Eq. (2.3). These equations are referred to as the second order Yang–Mills field equations:

$$D^m v_{mn} = 0, \tag{2.14}$$

and they consist of non-linear second-order partial differential equations. Self-dual and anti-self-dual gauge fields satisfy the second order Yang–Mills field equations via the Bianchi identity:

$$D^{m*} v_{mn} = 0. \tag{2.15}$$

Solutions of the second order Yang–Mills equations are however not necessarily self-dual or anti-self-dual.

Instantons have an action given by:

$$S_{\text{inst}} = -\frac{8\pi^2|k|}{g^2} - i\vartheta k = \begin{cases} -2\pi i k\tau & k > 0 \\ -2\pi i k\tau^* & k < 0 \end{cases}, \tag{2.16}$$

where τ is the complexified gauge coupling constant, given by:

$$\tau = \frac{4\pi i}{g^2} + \frac{\vartheta}{2\pi}, \tag{2.17}$$

which involves the classical gauge coupling g, as we treat only classical Yang–Mills gauge field theory for the purposes of this chapter.

In Subsection 2.2.1 below, we describe the first non-trivial solution discovered for the self-dual Yang–Mills field equations, and thus the first self-dual gauge field found, known as the BPST instanton.

2.2.1 The BPST Instanton

The first instanton solution of Yang–Mills gauge field theory was discovered by Belavin, Polyakov, Schvarts and Tyupkin (BPST) [1]. The BPST instanton was derived for four dimensional Yang–Mills gauge field theory with gauge group $SU(2)$. Originally termed 'pseudoparticles' by Belavin et. al [1], the instanton solution was given its modern name by 't Hooft [15], which describes the fact that the solution is localized in space and time. Viewed in five dimensional spacetime, the instanton would appear to be a localized particle of finite extent, solitonic in nature.

Anticipating the ADHM construction, which we shall describe in Section 2.3, we now introduce a quaternionic formalism for four dimensional Euclidean spacetime co-ordinates. This formalism will prove to be convenient when using the ADHM construction.

The Lorentz group in four dimensional Euclidean spacetime is $SO(4)$; this has a covering group $SU(2)_R \times SU(2)_L$. Then a spacetime 4-vector x_n, $n = 0, \ldots, 3$, can be written in the $(\mathbf{2}, \mathbf{2})$ representation of this product group, denoted in component form as the quaternions $x_{\alpha\dot{\alpha}}$ or $\bar{x}^{\dot{\alpha}\alpha}$. In this notation, $\alpha, \dot{\alpha} = 1, 2$ are Weyl spinor indices of the respective $SU(2)_L$ and $SU(2)_R$ groups. The explicit form of the quaternionic spacetime co-ordinates is given by:

$$x_{\alpha\dot{\alpha}} = x_n \sigma_{n\alpha\dot{\alpha}}, \quad \bar{x}^{\dot{\alpha}\alpha} = x_n \bar{\sigma}_n^{\dot{\alpha}\alpha}, \tag{2.18}$$

where $\sigma_{n\alpha\dot{\alpha}}$ and its Hermitian conjugate $\bar{\sigma}_n^{\dot{\alpha}\alpha}$ are four 2×2 matrices with components given by:

$$\sigma_n = (i\vec{\tau}, 1_{[2]\times[2]}), \quad \bar{\sigma}_n \equiv \sigma_n^\dagger = (-i\vec{\tau}, 1_{[2]\times[2]}), \tag{2.19}$$

in which $1_{[2]\times[2]}$ is the 2×2 identity matrix and τ^c, $c = 1, 2, 3$ are the three standard Euclidean Pauli matrices given in Appendix A.

The Weyl spinor indices $\alpha, \dot{\alpha}$ may be lowered and raised using the anti-symmetric tensors $\epsilon^{\alpha\beta}$ and $\epsilon_{\dot{\alpha}\dot{\beta}}$ whose form and action are defined in Appendix A.

One can write the quaternionic spacetime co-ordinates in a yet more explicit form, making use of the Euclidean Pauli matrices in Appendix A:

$$x_{\alpha\dot{\alpha}} = \begin{pmatrix} x_4 + ix_3 & x_2 + ix_1 \\ -x_2 + ix_1 & x_4 - ix_3 \end{pmatrix}, \quad \bar{x}^{\dot{\alpha}\alpha} = \begin{pmatrix} x_4 - ix_3 & -x_2 - ix_1 \\ x_2 - ix_1 & x_4 + ix_3 \end{pmatrix}. \tag{2.20}$$

Derivatives with respect to Euclidean spacetime can also be expressed in this quaternionic formalism, as:

$$\partial_{\alpha\dot{\alpha}} = \sigma_{n\alpha\dot{\alpha}} \partial_n, \quad \partial^{\dot{\alpha}\alpha} = \bar{\sigma}_n^{\dot{\alpha}\alpha} \partial_n. \tag{2.21}$$

Note that the derivatives in Eq. (2.21) do not give differentiation with respect to the quaternionic form of the spacetime co-ordinates, that is, $\partial/\partial x_{\alpha\dot{\alpha}}$.

Also required are the following Lorentz generators, which are combinations of the matrices in Eq. (2.19):

$$\sigma_{mn} = \tfrac{1}{4}(\sigma_m\bar{\sigma}_n - \sigma_n\bar{\sigma}_m), \quad \bar{\sigma}_{mn} = \tfrac{1}{4}(\bar{\sigma}_m\sigma_n - \bar{\sigma}_n\sigma_m). \tag{2.22}$$

These combinations are self-dual and anti-self-dual, respectively:

$$\sigma_{mn} = \tfrac{1}{2}\epsilon_{mnkl}\sigma_{kl}, \quad \bar{\sigma}_{mn} = -\tfrac{1}{2}\epsilon_{mnkl}\bar{\sigma}_{kl}. \tag{2.23}$$

We now turn to the BPST instanton. The BPST instanton is an instanton configuration of unit topological charge, referred to as a one-instanton, and is the unique one-instanton solution for Euclidean $SU(2)$ Yang–Mills gauge field theory. In general, instantons of topological charge k will be referred to as k-instantons. The BPST instanton gauge field has the explicit form:

$$v_m = \frac{2}{g}\frac{(x-X)_n\Omega^\dagger\sigma_{mn}\Omega}{(x-X)^2+\rho^2}. \tag{2.24}$$

The BPST instanton gauge field configuration has many important properties, some of which are generic to instantons, which we describe in a list below.

Properties of the BPST Instanton

The BPST instanton solution given in Eq. (2.24) has the following properties:

1. The instanton gauge field configuration is described by eight free, unconstrained parameters: one scale size (or dilatation) parameter ρ, which gives the size of the instanton; four spacetime co-ordinates contained in the 4-vector instanton position X_n, which represents the centre of the instanton; and three parameters contained in the global $SU(2)$ gauge rotations, or $SU(2)$ iso-orientations, $\Omega \in SU(2)$, of the instanton, which orient the instanton in $SU(2)$ group space. These are examples of instanton collective co-ordinates. The 'physical' degrees of freedom of the instanton are ρ and X_n, so that the BPST instanton can be interpreted as a five dimensional object [25]. More precisely, the $SU(2)$ one-instanton moduli space can be

interpreted as a five dimensional manifold, when global $SU(2)$ gauge rotations are excluded [10]. These parameters are associated with the classical global symmetries of the $SU(2)$ Yang–Mills gauge field theory which are broken by the instanton solution. These global symmetries are spacetime translations, scale transformations, and global gauge transformations.

2. The instanton solution possesses zero energy and hence has zero mass [62]. The gauge fields v_m described by the action $S[v_m]$ in Eq. (2.3) have no mass terms associated with them. More precisely, instanton gauge field configurations possess a vanishing energy-momentum tensor [15], which is a conserved quantity. For a Yang–Mills gauge theory, the energy-momentum tensor T_{mn} can be written as:

$$T_{mn} \;=\; \tfrac{1}{4}\left[(v_{ml} - {}^*v_{ml})(v_{nk} + {}^*v_{nk}) + (v_{nl} - {}^*v_{nl})(v_{mk} + {}^*v_{mk})\right]. \qquad (2.25)$$

For self-dual or anti-self-dual gauge fields, for which $v_{mn} = \pm^* v_{mn}$, from Eq. (2.25) it can be seen that instantons and anti-instantons possess a vanishing energy-momentum tensor, $T_{mn} = 0$ [15]. It is also notable that a vanishing Yang–Mills energy-momentum tensor implies that the Yang–Mills gauge field is self-dual or anti-self-dual; thus $T_{mn} = 0$ is a necessary and sufficient conditions for self-dual and anti-self-dual Yang–Mills gauge fields. Instantons can be considered as localized fluctuations of the pure massless gauge field vacuum. The result that $T_{mn} = 0$ is connected with this fact and the interpretation of instantons as trajectories in gauge field space which connect vacuum states. Other conserved quantities are also zero for instanton configurations; these include the isospin current and the dilatation current.

3. The instanton configuration breaks the product of the gauge group symmetry and the Lorentz group symmetry $SU(2) \times SU(2)_L$ to a diagonal subgroup symmetry. The $SU(2)$ group indices in Eq.(2.24) are identified with those of the $SU(2)_L$ subgroup.

4. Since the self-dual Yang–Mills field equations are locally gauge invariant, the BPST instanton in Eq. (2.24) exists in a particular local gauge. The particular gauge used in Eq. (2.24) is known as regular gauge. Local gauge transformations of Eq. (2.24)

correspond to differing values of the $SU(2)$ gauge group element Ω. Therefore the BPST instanton is equivalent to all other $SU(2)$ one-instantons with eight parameters, up to local gauge transformations.

5. The instanton gauge field has a gauge field strength given by:

$$v_{mn} = \frac{4}{g}\frac{\rho^2\Omega^\dagger\sigma_{mn}\Omega}{\left[(x-X)^2+\rho^2\right]^2}, \tag{2.26}$$

which is manifestly self-dual due to the self-duality of the generator σ_{mn}.

6. The BPST instanton is the most general one-instanton configuration for the gauge group $SU(2)$. For gauge group $SU(2)$, it has been shown that the most general k-instanton solution will have $8k$ free parameters [26, 54]. The BPST instanton has $8\times 1 = 8$ parameters and is thus the most general one-instanton solution for the self-dual Yang–Mills $SU(2)$ field equations.

The BPST anti-instanton, which has topological charge $k = -1$, and the gauge field strength of the BPST anti-instanton, may be obtained from Eqs. (2.24,2.26) by substituting the self-dual tensor $\sigma_{\alpha\dot{\alpha}}$ with the anti-self-dual tensor $\bar{\sigma}^{\dot{\alpha}\alpha}$ in these expressions. The particular local gauge in which the BPST instanton in Eq. (2.24) has been expressed is regular gauge. This is because in this gauge the BPST instanton is non-singular when $x_n = X_n$. At large distances, the BPST gauge field v_m has the asymptotic behaviour $v_m \sim 1/x$. This is an inconvenient gauge for semi-classical calculations involving instantons, upon which we will elaborate in Section 6.2 of Chapter 6. It is inconvenient as this behaviour leads to difficulties in constructing square-integrable quantities for the instanton gauge field. This technical difficulty can be resolved by using the BPST instanton, and instantons in general, in singular gauge. The BPST instanton in singular gauge can be obtained by a local gauge transformation of the regular gauge BPST instanton in Eq. (2.24), via:

$$v_m \to v'_m = Uv_mU^\dagger + \frac{i}{g}U\partial_m U^\dagger, \tag{2.27}$$

where $U(x)$ is the singular matrix:

$$U(x) = \frac{\Omega^\dagger\bar{\sigma}_m(x-X)_m}{|x-X|}. \tag{2.28}$$

The gauge transformation Eq. (2.27) using the singular matrix Eq. (2.28) is strictly only valid if the singular point $x_n = X_n$ is excluded from Euclidean spacetime $\mathbb{R}^4$. This apparent singularity of the instanton solution was introduced by a local gauge transformation Eq. (2.27), and so can also be removed via a local gauge transformation. A rigorous method for treating singular gauge transformations uses a punctured Euclidean spacetime with the point $x_n = X_n$ excluded. On such a spacetime, the singular gauge BPST instanton is regular. The long distance behaviour of the matrix $U(x)$ in Eq. (2.28) represents a bijective mapping from S^3_∞ to the gauge group $SU(2)$. It follows that the BPST instanton has unit topological charge, $k = 1$.

Furthermore, the topological charge of the BPST instanton in singular gauge is localized on the infinitesimal 3-sphere about the singular point $x_n = X_n$. This is unlike the topological charge of the regular gauge BPST instanton, which is localized on the 3-sphere at infinity, S^3_∞.

The result of the gauge transformation Eq. (2.27) is the BPST instanton in singular gauge, which reads:

$$v_m = \frac{2}{g}\frac{\rho^2(x-X)_n\Omega^\dagger\bar{\sigma}_{mn}\Omega}{(x-X)^2\left[(x-X)+\rho^2\right]}. \tag{2.29}$$

In singular gauge, the BPST instanton has its $SU(2)$ gauge group indices identified with the indices of the $SU(2)_R$ subgroup, unlike the regular gauge BPST instanton. At large distances, the singular gauge BPST instanton Eq. (2.29) has behaviour $v_m \sim 1/x^3$, which is useful for ensuring the convergence of integrals in instanton calculus. We shall describe instanton calculus in Chapter 6.

In Subsection 2.2.2 below, we describe the collective co-ordinates of generic instantons and their moduli space.

2.2.2 The Moduli Space of Instantons

In this subsection we describe the moduli space and collective co-ordinates of instantons. The moduli space is a key concept in the modern treatment of instantons and solitons in general. In Subsection 2.2.1, we described instantons in $SU(N)$ (or $U(N)$) Yang–Mills gauge field theory; here we continue with this choice of gauge group for the instanton moduli space. The moduli space of instantons is the space of gauge inequivalent solutions of the self-dual Yang–Mills field equations Eq. (2.11). The gauge inequivalence in this case is inequivalence up to local gauge transformations. Solutions which are related by a global gauge transformation are taken to be equivalent, as was stated for the BPST instanton in Subsection 2.2.1. Global and local gauge transformations affect the instanton solution differently, primarily because ordinary covariant gauge fixing conditions do not fix global gauge transformations [63]. Hence global gauge orbits must still be integrated over in semi-classical path integrals involving instantons.

The moduli space of $U(N)$ instantons can be described implicitly by the ADHM construction of instantons, which we describe in Section 2.3. Throughout this thesis, we use the term 'moduli space of instantons' to refer to the so-called extended moduli space of instantons. This is the moduli space of instantons which includes global gauge transformations as physical parameters, which we denote $\mathfrak{M}_k$. In mathematics literature, the usual moduli space of instantons, which has fewer parameters than $\mathfrak{M}_k$ as it excludes global gauge transformations, is often used. For use in the modern form of instanton calculus, however, the extended instanton moduli space is the appropriate space to integrate over. We shall return to this point again in Chapter 6.

The Moduli Space of $U(N)$ Instantons

As was described in Subsection 2.2, classical finite action solutions which minimize the four dimensional Euclidean Yang–Mills action can be classified according to the topological charge $k \in \mathbb{Z}$, also known as the instanton number. This integer labels the equivalence classes, or sectors, of instanton solutions which possess differing values of k. It follows that the complete moduli space of instantons must also be divided up into sectors labelled by k in the same way. We denote the moduli space of $SU(N)$ instantons of topological charge k as $\mathfrak{M}_k$. The moduli space of $U(N)$ k-instantons is also given by $\mathfrak{M}_k$. This is because the Abelian factor $U(1)$ in the isomorphism $U(N) \simeq SU(N) \times U(1)$ does not affect instanton solutions in commutative Euclidean spacetime. Hereon in this section we shall focus on $U(N)$ instantons.

The moduli space $\mathfrak{M}_k$ as a mathematical object is a manifold endowed with singularities, so that in the strictest sense, $\mathfrak{M}_k$ is actually a special kind of space. The singularities on

$\mathfrak{M}_k$ occur where instantons have zero scale size. The appearance of these singularities can be seen explicitly from the regular gauge BPST instanton Eq. (2.24), for which $\rho = 0$ allows the singular points $x_n = X_n$ to occur. We now describe the co-ordinates on the moduli space $\mathfrak{M}_k$.

Collective Co-ordinates

The instanton moduli space $\mathfrak{M}_k$ is a manifold with singularities, and can be assigned a local co-ordinate system to specify points in $\mathfrak{M}_k$. Note that a global co-ordinate system would not hold due to the presence of the singularities in $\mathfrak{M}_k$. The local co-ordinates on $\mathfrak{M}_k$ specify collective characteristics of self-dual gauge fields, and are termed collective co-ordinates. A general k-instanton gauge field, which we denote $v_m(x; \mathfrak{X}_\mu)$, will depend upon the Euclidean spacetime co-ordinates $x_n \in \mathbb{R}^4$ and the set of collective co-ordinates $\mathfrak{X}_\mu$, $\mu = 1, \ldots, \dim \mathfrak{M}_k$, where $\dim \mathfrak{M}_k$ is the dimension of the moduli space $\mathfrak{M}_k$. A generic instanton configuration is described by the set of collective co-ordinates $\{\mathfrak{X}_\mu\} = \{(X_n)_i, \rho_k\}$, $i = 1, \ldots, k$, where $\{(X_n)_i\}$ are the co-ordinates which specify the centre of the instanton and $\{\rho_k\}$ are the set of scale sizes which give the size of the instanton at various points on $\mathfrak{M}_k$. The physical interpretation of these particular collective co-ordinates is not always apparent, however.

For any given instanton solution, which will be localized on $\mathbb{R}^4$, there will exist a centre for the configuration. The instanton configuration can always be translated in $\mathbb{R}^4$, so that the collective co-ordinates $\{(X_n)_i\}$ can always be set to zero, centering the instanton at the spacetime origin. Since the co-ordinates $\{(X_n)_i\}$ are collective, the instanton gauge field cannot depend on $\{(X_n)_i\}$ in a way independent of the true physical spacetime co-ordinates x_n. Thus the instanton gauge field can only depend on the relative spacetime co-ordinates $(x_n - (X_n)_i)$. As aforementioned, this is because the instanton centre can always be translated to $\{(X_n)_i\} = (0,0,0,0)_i$ on $\mathfrak{M}_k$, and this must not affect the physical properties of the gauge field. This leads one to define the useful notion of the centred moduli space of instantons, which we write as $\widehat{\mathfrak{M}}_k$. On $\widehat{\mathfrak{M}}_k$, the centre of the instanton (given by $\{(X_n)_i\}$) has been factored out as:

$$\mathfrak{M}_k \;=\; \mathbb{R}^4 \times \widehat{\mathfrak{M}}_k. \tag{2.30}$$

The collective co-ordinates describing the instanton are associated with all of the gauge theory symmetries broken by the instanton solution. A generic instanton solution will break the translation invariance of the classical Euclidean Yang–Mills gauge theory and from this the collective co-ordinates $(X_n)_i$ originate. The internal space, or subspace, formed by the collective co-ordinates within the moduli space can be generated by the broken symmetries of the gauge theory acting on the instanton configuration. In the case of the translational collective co-ordinates $(X_n)_i$, the symmetries are spacetime translations, which can be generated by $\delta x(X) = -X_n \partial/\partial x_n$. Then one has:

$$v_m(x; X, \ldots) \;=\; e^{\delta x(X)} v_m(x; 0, \ldots) \;=\; v_m(x - X; 0, \ldots). \tag{2.31}$$

Note that not all collective co-ordinates originate from broken gauge theory symmetries, and that not all symmetries of the gauge theory lead to inequivalent collective co-ordinates. The symmetries of the Euler–Lagrange equations of the gauge theory exist also as symmetries on the moduli space $\mathfrak{M}_k$, but these symmetries act differently upon $\mathfrak{M}_k$. Some symmetries may leave $\mathfrak{M}_k$ invariant, others will be degenerate and map out the same subspace, whereas others will not be related to any symmetry.

In general, all collective co-ordinates are associated with zero modes of the gauge field v_m. Zero modes are physical fluctuations of the gauge field about the instanton solution which do not alter the value of the action $S[v_m]$; the term 'zero mode' derives from 'zero energy mode'. Hence such modes have zero action, and so do not contribute to the total action. To be more precise, it can shown that zero modes are suitably gauge fixed derivatives of the gauge field with respect to the collective co-ordinates of the theory:

$$\delta_\mu v_m \;=\; \partial v_m / \partial \mathfrak{X}_\mu. \tag{2.32}$$

The derivatives in Eq. (2.32) are, upon satisfying a certain gauge condition, then the zero modes corresponding to the collective co-ordinates $\mathfrak{X}_\mu$. Due to Eq. (2.32), one expects that the number of zero modes is equal to the dimension of the space of collective co-ordinates $\mathfrak{X}_\mu$, which is the instanton moduli space $\mathfrak{M}_k$, and below we outline the proof that this is indeed the case. We shall describe zero modes further in the context of the semi-classical approximation in Chapter 6.

Symmetries of the $U(N)$ Moduli Space

The moduli space $\mathfrak{M}_k$ has symmetries related to those of the classical Yang–Mills gauge theory. As in Yang–Mills gauge theory, the spacetime symmetries of $\mathfrak{M}_k$ are enhanced. In addition to the usual Poincaré spacetime symmetry of the four dimensional Euclidean spacetime, the gauge theory is also invariant under conformal transformation. Together, these symmetries, which comprise four translations, six rotations and four conformal translations, form a larger symmetry group than the spacetime Poincaré group. The conformal subgroup acts upon the quaternionic spacetime co-ordinate $x = x_{\alpha\dot{\alpha}} = x_n \sigma_{n\alpha\dot{\alpha}}$ as:

$$x \rightarrow x' = (Ax+B)(Cx+D)^{-1}, \quad AD - BC = 1, \tag{2.33}$$

where $\{A, B, C, D\} \in \mathbb{H}$ are quaternions. There are fifteen parameters in the conformal transformation Eq. (2.33); there are also fifteen generators of the enlarged spacetime symmetry group, so that the dimension of the conformal group is also fifteen. Since the moduli space describes gauge fields, $\mathfrak{M}_k$ is also invariant under global $U(N)$ gauge transformations.

Mathematical Structure of the Moduli Space

To conclude this subsection, we briefly describe the nature of the instanton moduli space $\mathfrak{M}_k$ in mathematical terms. The mathematical structure of $\mathfrak{M}_k$ is complicated, and we no not claim any rigour or completeness in what follows.

The dimension of $\mathfrak{M}_k$ can be deduced by applying the Atiyah–Singer index theorem [12] for elliptic operators [54]. This is done by considering the zero modes of the gauge fields described by the moduli space at a point on $\mathfrak{M}_k$. If v_m is an instanton solution of the self-dual Yang–Mills field equations in Eq. (2.11), let δv_m be some infinitesimal (quantum) fluctuation about v_m. Then to first order in the fluctuations δv_m, the self-dual Yang–Mills equations for the gauge field $v_m + \delta v_m$ is given by [40, 54]:

$$\mathcal{D}_m \delta v_n - \mathcal{D}_n \delta v_m = \epsilon_{mnkl} \mathcal{D}_k \delta v_l. \tag{2.34}$$

Equation (2.34) can be rewritten in terms of the quaternionic formalism, developed above for the spacetime co-ordinate x_n, applied to the gauge field v_m, as:

$$\vec{\tau}^{\dot{\alpha}}_{\dot{\beta}}\bar{\sigma}^{\dot{\beta}\alpha}_{n}\mathcal{D}\delta v_{\alpha\dot{\alpha}} \; = \; 0, \tag{2.35}$$

which represents three independent equations. As previously, $\vec{\tau}$ are the three standard Euclidean Pauli matrices, given in Appendix A. Upon imposing the gauge condition necessary to remove local gauge transformations from Eq. (2.35), which requires that zero modes are orthogonal to gauge transformations, the zero mode condition Eq. (2.35) can be expressed as a single quaternionic equation:

$$\bar{\sigma}^{n\dot{\alpha}\alpha}\mathcal{D}_n\delta v_{\alpha\dot{\beta}} \; = \; 0, \tag{2.36}$$

which is the covariant Weyl equation for a Weyl spinor $\delta v_{\alpha\dot{\beta}}$. The free index $\dot{\beta}$ indicates that there are two independent solutions of this equation for each gauge field zero mode. Since $\mathcal{D}_n$ is an elliptic operator, the Atiyah–Singer index theorem can be applied to it. In a lengthy calculation involving algebraic geometry, the dimension of the instanton moduli space $\mathfrak{M}_k$ can be determined from this equation by counting the number of gauge field zero modes [12, 26]. For the gauge group $U(N)$ or $SU(N)$, the result is that the dimension of the instanton moduli space is:

$$\dim \mathfrak{M}_k \; = \; 4Nk. \tag{2.37}$$

This general result gives the number of collective co-ordinates which the instantons described by the $U(N)$ k-instanton moduli space $\mathfrak{M}_k$ possess. Hence the number of free, unconstrained parameters which completely describe the most general exact $U(N)$ k-instanton gauge field configuration is $4Nk$. One can verify this result for a special case: for the BPST instanton, Eq. (2.24), one has $N = 2$ and $k = 1$, which agrees with Eq. (2.37).

In general terms, the instanton moduli space $\mathfrak{M}_k$ is a non-compact complex manifold with special properties. It is a hyper-Kähler space with singularities of conical type [67]. This is related to the fact that Euclidean spacetime is a hyper-Kähler space. A proof that $\mathfrak{M}_k$ is a hyper-Kähler space has been given in [68]. Furthermore, $\mathfrak{M}_k$ possesses sets of inequivalent complex structures, which correspond to the sectors of different topological charge. The instanton moduli space can also be considered a Riemannian manifold with a metric defined as an inner product of gauge field zero modes:

$$g_{\mu\nu}(X) = -2g^2 \int d^4x \, \mathrm{tr}_N \, \delta_\mu v_n \delta_\nu v_n, \tag{2.38}$$

where the gauge field zero modes $\delta_\mu v_n$ have had the appropriate gauge fixing condition imposed on them, as defined in Eq.(2.32). This metric shall appear again in Chapter 6, where the volume form for integration on $\mathfrak{M}_k$, using the collective co-ordinate method, will depend upon $g_{\mu\nu}(X)$.

The instanton moduli space can be derived using a method known as the hyper-Kähler quotient construction [67], which we shall mention briefly in the next section, Section 2.3. In Section 2.3 we describe the ADHM construction, a method by which all instantons, that is, all self-dual (and anti-self-dual) gauge fields with arbitrary classical gauge group and topological charge k, can be constructed.

2.3 The ADHM Construction of Instantons

In this section we describe the method for constructing instantons discovered by Atiyah, Drinfeld, Hitchin and Manin (ADHM) [8]. The ADHM construction a method which gives the general solution to the self-dual and anti-self-dual Yang–Mills field equations for all values of the topological charge k and for all classical gauge groups, namely the groups $U(N)$, $O(N)$ and their special forms $SU(N)$ and $SO(N)$, and also $Sp(N)$. We note that the ADHM construction is not the only method which solves the self-dual Yang–Mills equations, but it is the most widely known and used; also, it is the most successful method for constructing general multi-instanton solutions. The ADHM construction is a remarkable mathematical achievement. Through sophisticated algebraic geometry, the construction reduces the self-dual Yang–Mills field equations, which are first order non-linear partial differential equations, to non-linear algebraic equations which can be written in terms of matrices. Essentially, the ADHM construction provides a way to describe and parameterize the moduli space of instantons.

However, the construction does contain some shortcomings. The single most problematic aspect of the ADHM construction is that the self-dual gauge fields resulting from it are defined only implicitly, in terms of a set of non-trivial constraints known as the ADHM constraints. Another problem, which presents less difficulty, is the amount of redundancy which it is necessary to remove from the solutions it generates. This redundancy must be removed if ADHM instanton configurations are to be used in semi-classical calculations. In Subsection 2.3.1 we will focus on the ADHM construction for $U(N)$ instantons, which are described by the moduli space $\mathfrak{M}_k$ introduced in Section 2.2, and in particular we describe the explicit construction of the exact general $U(N)$ one-instanton solution. This is followed in Subsection 2.3.2 by a derivation of the exact general $U(N)$ ADHM two-instanton, the first known multi-instanton for the gauge group $U(N)$. The $U(N)$ ADHM two-instanton was determined by solving the ADHM constraints for this case. We also describe the ADHM construction for the $U(N)$ three-instanton. In Subsection 2.3.3 we consider ADHM multi-instantons with gauge group $Sp(N)$. These instanton configurations are useful through the isomorphism $Sp(1) \simeq SU(2)$, which provide a more simple and efficient parameterization of $SU(2)$ multi-instantons than the $U(N)$ formalism. In particular, we describe the existing solutions for the $Sp(N)$ ADHM three-instanton and describe attempts to solve the general $Sp(N)$ three-instanton constraints.

In this section we make use of the reviews in [49, 50, 97] and the reviews on the modern treatment of the ADHM construction contained in [223, 224, 225]. We also refer to the original papers concerning the ADHM construction in [8, 22, 33, 34], and note the seminal works which the ADHM construction is based upon [3, 4, 5, 6, 7, 9, 10, 35]. Also of note are the mathematical works [52, 53, 89] concerning the ADHM construction, and the reviews of its mathematical origins [13, 14, 51, 91, 92, 93, 94]. We also refer to the works regarding the properties and symmetries of instantons in [18, 19, 25, 26, 27, 28, 29], the reviews [63, 64] and the related works [20, 21, 30, 31, 32, 77]. Other works concerning ADHM multi-instantons in [22, 23, 24, 36, 101] are also referred to in later subsections.

2.3.1 $U(N)$ ADHM Instantons

We begin with a brief exposition of the ADHM construction for the gauge group $U(N)$, closely following the formalism developed in [223, 224, 225]. Unlike the formalism of [223], we shall work in Euclidean spacetime as is conventional for instantons. The construction is described in terms of an ansatz involving the 'pure gauge' condition Eq. (2.7). The fundamental matrices which are used in the ADHM construction are rectangular matrices whose elements are complex parameters, which in this treatment can also be combined into quaternions. We describe the properties of quaternions in Appendix B. We note that it is not necessary to use quaternions for the $U(N)$ ADHM construction, but it provides a convenient unified treatment for the $Sp(N)$ and $O(N)$ ADHM constructions also. A version of the $U(N)$ ADHM construction which does not use quaternions, but which uses complex variables instead, is given in [33].

The notation and conventions used for our treatment of the $U(N)$ ADHM construction are as follows. The gauge field $v_m(x)$ is an $N \times N$ anti-Hermitian matrix of complex elements, and is a function of the spacetime co-ordinate x. Anti-Hermicity of v_m, which conventionally is a Hermitian matrix (since it is observable), is achieved by setting $v_m \to iv_m$ and also $v_{mn} \to iv_{mn}$. The anti-Hermicity of v_m is in fact already built into the ADHM construction. Factors of the gauge coupling g are kept explicit, as is done in the other chapters.

As in [223], in our notation an over-bar indicates Hermitian conjugation for matrix quantities, and complex conjugation for scalar quantities; thus $\bar{A} \equiv A^\dagger$ for a matrix A, and $\bar{b} \equiv b^*$ for a scalar b. However, we also employ the asterisk for the complex conjugation of matrices. Multiplication of quaternionic matrices follows the conventions of Appendix B. The modulus of quaternions and complex matrices also follows these conventions, and is such that for $A \in \mathbb{H}$, $A\bar{A} = \bar{A}A = |A|^2$. Matrix dimensions are given explicitly as encased subscripts; $M_{[a]\times[b]}$ indicates that the matrix M has a rows and b columns. Matrix multiplication is written in terms of these indices as $(AB)_{[a]\times[c]} = A_{[a]\times\underline{[b]}}B_{\underline{[b]}\times[c]}$, where the underlining of the indices indicates the contraction of matrix indices.

In the ADHM construction for $U(N)$, one begins with an $(N+2k) \times 2k$ complex matrix $\Delta_{[N+2k]\times[2k]}$, which is defined to be linear in the quaternionic spacetime co-ordinate x, defined in Eq. (2.18):

$$\Delta(x) \equiv \Delta_{[N+2k]\times[2k]}(x) = a_{[N+2k]\times[2k]} + b_{[N+2k]\times[k]\times\underline{[2]}}x_{\underline{[2]}\times[2]}. \tag{2.39}$$

The index $[2k]$ has been decomposed as the direct product of indices $[k] \times [2]$ in order to exhibit the contraction of indices in the matrix multiplication. We shall refer to this decomposition of matrix indices as the 'ADHM index convention.' The matrices a and b are complex-valued constant matrices which contain the ADHM data describing the instanton, and comprise an overcomplete set of k-instanton collective co-ordinates.

The nullspace of the Hermitian conjugate matrix $\bar{\Delta}(x)$ is an N-dimensional space, which has basis vectors that form an $(N+2k)\times N$-dimensional complex matrix $U(x)$, where:

$$\bar{\Delta}_{[2k]\times\underline{[N+2k]}}U_{\underline{[N+2k]}\times[N]} = \bar{U}_{[N]\times\underline{[N+2k]}}\Delta_{\underline{[N+2k]}\times[2k]} = 0. \tag{2.40}$$

The matrix $U(x)$ is orthonormalized to the $N\times N$ unit matrix:

$$\bar{U}_{[N]\times\underline{[N+2k]}}U_{\underline{[N+2k]}\times[N]} = 1_{[N]\times[N]}. \tag{2.41}$$

The instanton gauge field $v_m(x)$ can be constructed from the matrix $U(x)$. When the topological charge is zero, $k=0$, the gauge field is given by a gauge transformation of the vacuum ('pure gauge'):

$$v_{m[N]\times[N]} = \frac{1}{g}\bar{U}_{[N]\times\underline{[N+2k]}}\partial_m U_{\underline{[N+2k]}\times[N]}, \tag{2.42}$$

which automatically satisfies the self-dual Yang–Mills field equations (Eq. (2.11)). Equation (2.42) is identical to Eq. (2.7), when written in matrix form with our convention of an anti-Hermitian gauge field, which means the factor of i in Eq. (2.7) does not appear in Eq. (2.42). In the ADHM construction, the condition Eq. (2.42) is taken to give a solution to the self-dual Yang–Mills field equations for all non-zero values of k. This is the central ansatz of the ADHM construction. The ADHM ansatz implies the following factorization condition for $\Delta(x)$:

$$\bar{\Delta}_{[2]\times[k]\times\underline{[N+2k]}}\Delta_{\underline{[N+2k]}\times[k]\times[2]} = 1_{[2]\times[2]}f^{-1}_{[k]\times[k]}, \tag{2.43}$$

where $f(x)$ is an arbitrary x-dependent $k\times k$-dimensional Hermitian matrix.

When combined with the nullspace condition in Eq. (2.40), Eq. (2.43) then implies the completeness relation, which is required for consistency:

$$\Delta_{[N+2k]\times\underline{[k]}\times\underline{[2]}}f_{\underline{[k]}\times\underline{[k]}}\bar{\Delta}_{\underline{[2]}\times\underline{[k]}\times[N+2k]} = 1_{[N+2k]\times[N+2k]} - U_{[N+2k]\times\underline{[N]}}\bar{U}_{\underline{[N]}\times[N+2k]}. \tag{2.44}$$

The relation Eq. (2.44) also provides a test of the validity of the ADHM ansatz stated above.

Using Eqs. (2.42,2.43,2.44) with integration by parts, and using the short-hand notation

$X_{[m}Y_{n]} = X_mY_n - X_nY_m$, the gauge field strength can then be expressed as:

$$\begin{aligned} v_{mn} &\equiv \partial_{[m}v_{n]} + gv_{[m}v_{n]} \;=\; \partial_m v_n - \partial_n v_m + g[v_m, v_n] \\ &= g^{-1}\partial_{[m}(\bar{U}\partial_{n]}U) + g^{-1}(\bar{U}\partial_{[m}U)(\bar{U}\partial_{n]}U) \;=\; g^{-1}\partial_{[m}\bar{U}(1-U\bar{U})\partial_{n]}U \\ &= g^{-1}\partial_{[m}\bar{U}\Delta f\bar{\Delta}\partial_{n]}U \;=\; g^{-1}\bar{U}\cdot\partial_{[m}\Delta f\partial_{n]}\bar{\Delta}\cdot U \\ &= g^{-1}\bar{U}b\sigma_{[m}\bar{\sigma}_{n]}f\bar{b}U \;=\; 4g^{-1}\bar{U}b\sigma_{mn}f\bar{b}U, \end{aligned} \tag{2.45}$$

where σ_{mn} is the self-dual numerical tensor defined in Eq. (2.19). Since σ_{mn} is manifestly self-dual, it follows that the field strength v_{mn} is also self-dual. Hence the ansatz is correct, and it can be shown that this construction gives all self-dual $U(N)$ gauge fields of arbitrary topological charge k, and thus gives the general solution of the self-dual $U(N)$ Yang–Mills field equations. This construction can also be adapted for the general solution of the anti-self-dual Yang–Mills field equations, in which the gauge field strength is manifestly anti-self-dual due to a factor of the anti-self-dual tensor $\bar{\sigma}_{mn}$ in the same position as σ_{mn} has in Eq. (2.45).

The instanton gauge field so constructed has gauge group $U(N)$. To specify the $SU(N)$ instanton gauge field, one can perform a global gauge transformation on the matrix U, given by $U \to Ug_1$, where $g_1 \in U(1)$. As stated previously in Subsection 2.2.2, instantons with gauge group $U(N)$ or $SU(N)$ are described by the same moduli space, denoted $\mathfrak{M}_k$. The $U(N)$ ADHM construction, which provides an implicit description of $\mathfrak{M}_k$, thus also describes $SU(N)$ instantons.

Continuing with the formalism of [223], we assign the following indices to the objects constituting the 'ADHM data' (the matrices U, Δ, a, b and f, which involve the matrices σ and x):

$$\begin{aligned} \text{Instanton number indices } [k] &: \; 1 \le i, j, l \cdots \le k \\ \text{Gauge group indices } [N] &: \; 1 \le u, v \cdots \le N \\ \text{ADHM indices } [N+2k] &: \; 1 \le \lambda, \mu \cdots \le N+2k \\ \text{Quaternionic (Weyl) indices } [2] &: \; \alpha, \beta, \dot{\alpha}, \dot{\beta} \cdots = 1, 2 \\ \text{Lorentz indices } [4] &: \; m, n \cdots = 0, 1, 2, 3. \end{aligned} \tag{2.46}$$

No extra notation is required for the $2k$-dimensional column index attached to Δ, a and b, since it can be factored as $[2k] = [k] \times [2] = j\dot{\beta}$ according to the ADHM index convention. In these index conventions, the matrices comprising the ADHM data can be assigned explicit indices. The fundamental matrix $\Delta(x)$ defined in Eq. (2.39) can be written as:

$$\Delta_{\lambda i\dot{\alpha}}(x) = a_{\lambda i\dot{\alpha}} + b^{\alpha}_{\lambda i}x_{\alpha\dot{\alpha}}\,, \qquad \bar{\Delta}^{\dot{\alpha}\lambda}_{i}(x) \equiv (\Delta_{\lambda i\dot{\alpha}})^{*} = \bar{a}^{\dot{\alpha}\lambda}_{i} + \bar{x}^{\dot{\alpha}\alpha}\bar{b}^{\lambda}_{\alpha i}, \tag{2.47}$$

The columns of the matrix $\Delta(x)$ must remain linearly independent for all values of x in order to avoid singularities in the integrand in the definition of the topological charge k, Eq. (2.4) [33]. This is a non-degeneracy condition which can also be expressed as demanding that the mappings represented by Δ and $\bar{\Delta}$ satisfy certain conditions. These are that the mapping $\Delta_{\dot{\alpha}}(x) : \mathbb{C}^{k} \to \mathbb{C}^{N+2k}$ is injective and that the mapping $\bar{\Delta}^{\dot{\alpha}}(x)$ is surjective. Hence the matrix $\Delta(x)$ is invertible.

The factorization condition in Eq. (2.43) can now be written as:

$$\bar{\Delta}^{\dot{\alpha}\lambda}_{i}\,\Delta_{\lambda j\dot{\beta}} = \delta^{\dot{\alpha}}_{\ \dot{\beta}}\,(f^{-1})_{ij}, \tag{2.48}$$

and the non-degeneracy condition stated above ensures that the inverse matrix f^{-1} exists. The nullspace condition Eq. (2.40) and the orthonormalization of the matrix U in Eq. (2.41) can also be expressed in terms of the index conventions above. The matrix U is then a $(N+2k)\times N$ dimensional complex matrix $U_{\lambda u}(x)$, where $u = 1,\ldots,N$. Equations (2.40,2.41) can then be written, respectively, as:

$$\bar{\Delta}^{\dot{\alpha}\lambda}_{i}U_{\lambda u} = 0 = \bar{U}^{\lambda u}\Delta_{\lambda i\dot{\alpha}}, \tag{2.49}$$

$$\bar{U}^{\lambda u}U_{\lambda v} = \delta_{uv}. \tag{2.50}$$

The equation which defines the gauge field in terms of U, Eq. (2.42) can also be re-expressed, as:

$$(v_m)_{uv} = \frac{1}{g}\bar{U}^{\lambda}_{u}\partial_m U_{\lambda v}, \tag{2.51}$$

which for $k = 0$ gives a gauge transformation of the gauge field vacuum $v_m = 0$ ('pure gauge'). The completeness relation Eq. (2.44) can also be used to define a projection operator, $\mathcal{P}$:

$$\mathcal{P}^{\mu}_{\lambda} \equiv U_{\lambda u}\bar{U}^{\mu}_{u} = \delta^{\mu}_{\lambda} - \Delta_{\lambda i\dot{\alpha}}f_{ij}\bar{\Delta}^{\dot{\alpha}\mu}_{j}, \tag{2.52}$$

where $\mathcal{P}$ satisfies the properties required of a projection operator due to the properties of U:

$$\mathcal{P}^{2} = \mathcal{P}, \ \bar{\mathcal{P}} = \mathcal{P}. \tag{2.53}$$

The definition of the projection operator $\mathcal{P}$ in Eq. (2.52) can also be written explicitly using matrix dimensions as:

$$\mathcal{P}_{[N+2k]\times[N+2k]} \equiv U_{[N+2k]\times\underline{[N]}}\bar{U}_{\underline{[N]}\times[N+2k]}, \tag{2.54}$$
$$= 1_{[N+2k]\times[N+2k]} - \Delta_{[N+2k]\times\underline{[k]}\times\underline{[2]}}f_{\underline{[k]}\times\underline{[k]}}\bar{\Delta}_{\underline{[2]}\times\underline{[k]}\times[N+2k]}. \tag{2.55}$$

One can verify that the winding number κ is equal to the instanton charge k of the solutions generated by the ADHM construction. The instanton number k of an ADHM configuration can be calculated in terms of ADHM data by using an identity first derived by Osborn [47]. This identity expresses the gauge group trace $\mathrm{tr}_N\, v_{mn}v^{mn}$ in the ADHM instanton background as:

$$-\frac{g^2}{16\pi^2}\int d^4\, x\mathrm{tr}_N v_{mn}v^{mn} = \frac{1}{16\pi^2}\int d^4x\, \Box^2\, \mathrm{tr}_N \log f. \tag{2.56}$$

From the factorization condition in Eq. (2.48), the matrix $f(x)$ has asymptotic behaviour at large $|x|$ such that:

$$\lim_{|x|\to\infty} f(x) = \frac{1}{x^2}1_{[k]\times[k]}. \tag{2.57}$$

Upon substituting Eq. (2.57) into Eq. (2.56), one deduces that the right hand side of Eq. (2.56) is equal to the winding number κ and it follows that the topological charge of an ADHM gauge field configuration is always equal to its instanton charge k.

We now turn to the ADHM constraints. The definition of $\Delta(x)$ in Eq. (2.39) and the factorization condition Eq. (2.48) imply a set of x-independent conditions on the matrices a and b, since $f_{ij}(x)$ is arbitrary, upon expanding $\Delta(x)$ as $\Delta(x) = a+bx$. These constraints have the form:

$$\bar{a}_i^{\dot{\alpha}\lambda}\, a_{\lambda j\dot{\beta}} = (\bar{a}a)_{ij}\delta^{\dot{\alpha}}_{\dot{\beta}} \propto \delta^{\dot{\alpha}}_{\dot{\beta}} \tag{2.58}$$
$$\bar{a}_i^{\dot{\alpha}\lambda}\, b^{\beta}_{\lambda j} = \bar{b}_i^{\beta\lambda}\, a^{\dot{\alpha}}_{\lambda j} \tag{2.59}$$
$$\bar{b}^{\lambda}_{\alpha i}\, b^{\beta}_{\lambda j} = (\bar{b}b)_{ij}\delta^{\beta}_{\alpha} \propto \delta^{\beta}_{\alpha}. \tag{2.60}$$

Together, the three conditions in Eqs. (2.58,2.59,2.60) constitute the $U(N)$ ADHM constraints in their original form. We refer to these constraints as the original ADHM constraints, since the constraints in Eqs.(2.58-2.60) contain an unfixed global $U(k)$ gauge symmetry which we describe below. Equations (2.59,2.60) also imply the relation:

$$\bar{\Delta}^{\dot{\alpha}}b^{\alpha} = \bar{b}^{\alpha}\Delta^{\dot{\alpha}}, \tag{2.61}$$

which can prove useful in manipulating ADHM matrices, often referred to as 'ADHM algebra'.

The matrices a and b contain the collective co-ordinates of the $U(N)$ k-instanton gauge field configuration. The number of instanton collective co-ordinates increases as k^2. However, the number of physical collective co-ordinates required to describe the $U(N)$ k-instanton moduli space is $4Nk$, as given in Eq. (2.37). This counting of parameters includes global gauge rotations of the gauge field [26]. Hence, taken together, the matrices a and b form an overcomplete set of collective co-ordinates. Some of the redundancy contained in a and b can be removed via the following x-independent transformations under which the $U(N)$ ADHM construction is invariant:

$$\begin{aligned} \Delta_{[N+2k]\times[k]\times[2]} &\to \Lambda_{[N+2k]\times\underline{[N+2k]}}\Delta_{\underline{[N+2k]}\times\underline{[k]}\times[2]}B^{-1}_{\underline{[k]}\times[k]}, \\ U_{[N+2k]\times[N]} &\to \Lambda_{[N+2k]\times\underline{[N+2k]}}U_{\underline{[N+2k]}\times[N]}, \\ f_{[k]\times[k]} &\to B_{[k]\times\underline{[k]}}f_{\underline{[k]}\times\underline{[k]}}B^{\dagger}_{\underline{[k]}\times[k]}, \end{aligned} \tag{2.62}$$

where $\Lambda \in U(N+2k)$ and $B \in GL(k,\mathbb{C})$.

One can use the symmetries in Eq. (2.62) to bring the representation of a and b to the canonical form given by Corrigan et. al [33]. The canonical form is obtained by removing the degrees of freedom contained in the matrix b, so that one has:

$$a_{[N+2k]\times[2k]} = \begin{pmatrix} w_{[N]\times[2k]} \\ a'_{[2k]\times[2k]} \end{pmatrix}, \qquad b_{[N+2k]\times[2k]} = \begin{pmatrix} 0_{[N]\times[2k]} \\ 1_{[2k]\times[2k]} \end{pmatrix}, \tag{2.63}$$

which implies that all of the physical degrees of freedom describing the instanton, namely the collective co-ordinates, reside in the matrix a.

We note that this use of the symmetries in Eq. (2.62) is not unique. It is possible that other ways of using the symmetries in Eq. (2.62) exist which simplify the $U(N)$ ADHM construction yet further. The canonical form in which we have written a and b in Eq. (2.63) can also be expressed making the matrix indices explicit. Using the ADHM index decomposition $\lambda = (u + i\alpha)$, one has, in the canonical form, for b:

$$b^{\beta}_{\lambda j} = b^{\beta}_{(u+i\alpha)j} = \begin{pmatrix} 0 \\ \delta^{\beta}_{\alpha}\delta_{ij} \end{pmatrix}, \quad \bar{b}^{\lambda}_{\beta j} = \bar{b}^{(u+i\alpha)}_{\beta j} = (0 \; \delta^{\beta}_{\alpha}\delta_{ji}), \tag{2.64}$$

and for a, written in a similar form as b:

$$a_{\lambda j\dot{\alpha}} = a_{(u+i\alpha)j\dot{\alpha}} = \begin{pmatrix} w_{uj\dot{\alpha}} \\ (a'_{\alpha\dot{\alpha}})_{ij} \end{pmatrix}, \quad \bar{a}_j^{\dot{\alpha}\lambda} = \bar{a}_j^{\dot{\alpha}(u+i\alpha)} = \left(\bar{w}_{ju}^{\dot{\alpha}} \ (\bar{a}'^{\dot{\alpha}\alpha})_{ji}\right). \tag{2.65}$$

The submatrix a' has elements $a' \equiv (a'_{\alpha\dot{\alpha}})_{ij}$ which can also be represented in a quaternionic basis:

$$(a'_{\alpha\dot{\alpha}})_{ij} = (a'_n)_{ij}\sigma_{n\alpha\dot{\alpha}}, \qquad (\bar{a}'^{\alpha\dot{\alpha}})_{ij} = (a'_n)_{ij}\bar{\sigma}_n^{\alpha\dot{\alpha}}, \tag{2.66}$$

which follows the quaternionic formalism used for the spacetime co-ordinate x and other quantities so far introduced. The canonical form of b also obeys the identity:

$$\bar{b}_\alpha b^\beta = \delta_\alpha^\beta 1_{[k]\times[k]}. \tag{2.67}$$

The inverse of the ADHM matrix $f = f_{ij}(x)$ can also be expressed in terms of the submatrices of a, as:

$$f^{-1} = 2\left(\bar{w}^{\dot{\alpha}}w_{\dot{\alpha}} + (a'_n + x_n 1_{[k]\times[k]})^2\right). \tag{2.68}$$

The submatrix w is a complex valued matrix which can also be represented as a quaternion using this formalism.

In addition to invariance under the transformations in Eq. (2.62) there exists an auxiliary, or residual, symmetry arising from the symmetry of the ADHM construction in Eq. (2.62). The canonical form of b given in Eq. (2.63) is invariant under global $U(k)$ rotations. These appear since $U(k)$ is a subgroup of the $U(N+2k)\times GL(k,\mathbb{C})$ symmetry which acts as in Eq. (2.62). The $U(k)$ symmetry group acts upon the matrix $\Delta(x) = \Delta_{[N+2k]\times[2k]}(x)$ as:

$$\Delta_{[N+2k]\times[2k]} \to \begin{pmatrix} 1_{[N]\times[N]} & 0_{[2k]\times[N]} \\ 0_{[N]\times[2k]} & \bar{\Lambda}_{[2k]\times[2k]} \end{pmatrix} \Delta_{[N+2k]\times[2k]}\Lambda_{[2k]\times[2k]}, \tag{2.69}$$

where $\Lambda_{[2k]\times[2k]} = \Omega_{[k]\times[k]}1_{[2]\times[2]}$ and $\Omega_{[k]\times[k]} \in U(k)$. This auxiliary $U(k)$ symmetry can be employed to simplify the final form of solutions of the ADHM constraints, or equivalently, the ADHM constraints themselves. The $U(k)$ residual symmetry acts as the matrix Λ in Eq. (2.62), and has the form:

$$\Lambda = \begin{pmatrix} 1_{[N]\times[N]} & 0 \\ 0 & \Omega 1_{[2]\times[2]} \end{pmatrix}, \quad \Omega \in U(k), \tag{2.70}$$

where $\Omega = B$ in Eq. (2.62) when Λ has the form given. The $U(k)$ symmetry acts as a non-trivial transformation on the submatrices of a, as:

$$w_{ui\dot{\alpha}} \rightarrow w_{\dot{\alpha}}\Omega, \;\; a'_n \; = \; \Omega^{\dagger}a'_n\Omega. \tag{2.71}$$

With a and b in the canonical form, the third original ADHM constraint Eq. (2.60) is automatically satisfied. The remaining original ADHM constraints then give the following $k \times k$ matrix equations:

$$\begin{aligned} \mathrm{tr}_2(\tau^{c\dot{\alpha}}_{\dot{\beta}}\bar{a}^{\dot{\beta}}a_{\dot{\alpha}}) &= \vec{\tau}^{\dot{\alpha}}_{\dot{\beta}}\left(\bar{w}^{\dot{\beta}}w_{\dot{\alpha}} + \bar{a}'^{\dot{\beta}\alpha}a'_{\alpha\dot{\alpha}}\right) = 0, & (2.72)\\ (a'_n)^{\dagger}_{ij} &= (a'_n)_{ij}, & (2.73) \end{aligned}$$

where the trace tr_2 in Eq. (2.72) is over the Weyl indices of the Pauli matrices τ^c, $c = 1, 2, 3$; the three Pauli matrices τ^a have been used to contract the product $(\bar{a}a)$. Hence there are three distinct equations in Eq. (2.72). The constraints in Eqs. (2.72,2.73) constitute the 'ADHM constraints' referred to previously and henceforth. The unknown variables in the construction are then the submatrices contained in the matrix $a_{\dot{\alpha}} = \{w_{\dot{\alpha}}, a'_n\}$, in which a'_n are assumed to be Hermitian matrices, as in Eq. (2.73). Note that Eq. (2.73) is actually a remnant of the formalism employed here, and can be discarded in alternative formalisms, such as that of Corrigan et. al [33]. It states simply that the respective real and imaginary parts of the complex matrices a'_n are Hermitian quantities, which must be the case when forming a complex matrix. Thus the non-trivial ADHM constraints are given by Eq. (2.72). Once the $U(k)$ residual symmetry has been fixed, from these constraints one can derive $U(N)$ instantons which possess no additional symmetries peculiar to the ADHM construction.

The ADHM constraints Eqs. (2.72,2.73) present both remarkable progress and considerable difficulty in the study of instantons. The dimensions of the matrices involved in the construction change with k, and so far the only ADHM instantons found have been obtained for particular values of k. Whilst the constraints implicitly define all k-instantons for the $U(N)$ Yang–Mills gauge field theory, extracting explicit instanton configurations from them has met with only limited success. In general, the ADHM constraints are a set of non-linear, non-trivial coupled simultaneous matrix equations. In terms of their components, they are a set of non-linear coupled simultaneous algebraic equations, at most quadratic and bilinear in the elements of the matrix a. Explicit general solutions of the ADHM constraints have so far only been found for $k = 1$ and $k = 2$. We describe both of these cases in detail below, and in Subsection 2.3.2, respectively. As k increases, the number of constraints to be solved increases approximately as $2k$, and the construction quickly becomes exceptionally difficult to extract solutions from. There is an order of magnitude increase in the complexity of the constraints as k increases by one unit. In Subsection 2.3.2 we shall describe the $U(N)$ $k = 2$ and $k = 3$ ADHM constraints, which clearly illustrates this increase. The consensus is that the ADHM constraints for any gauge group are in fact not possible to explicitly solve for $k \geq 4$.

When $N = 2$, one can adopt the $Sp(1) \simeq SU(2)$ ADHM constraints, which we also describe below. More progress has been made in solving the $Sp(N)$ constraints, since explicit solutions are known for $k \leq 3$. However, this is still very limited, and an explicit k-instanton solution does not appear possible for the ADHM constraints. It would also appear that there is insufficient information regarding the underlying principles involved in solving the ADHM constraints. If such principles do exist, an algorithm or iterative process for obtaining explicit instanton solutions could possibly be developed.

The explicit form of the ADHM matrix a can be written out using the quaternionic formalism which has been employed so far, and in our notation is given by:

$$a_{[N+2k]\times[2k]} = \begin{pmatrix} w_{[N]\times[2k]} \\ a'_{[2k]\times[2k]} \end{pmatrix} = \begin{pmatrix} w_{1[N]\times[k]} & w_{2[N]\times[k]} \\ (a'_4 + ia'_3)_{[k]\times[k]} & (a'_2 + ia'_1)_{[k]\times[k]} \\ (-a'_2 + ia'_1)_{[k]\times[k]} & (a'_4 - ia'_3)_{[k]\times[k]} \end{pmatrix} \tag{2.74}$$

We note that the ADHM matrix a in Eq. (2.74) is a matrix of complex submatrices which may be expressed in terms of quaternions (written as 2×2 complex matrices); however, we shall treat a as having complex submatrices. Hence the $U(N)$ k-instanton ADHM constraints in Eqs. (2.72,2.73) can be expressed in terms of the complex valued component matrices $(a'_4 + ia'_3)_{[k]\times[k]}$, $(a'_2 + ia'_1)_{[k]\times[k]}$, $w_{1[N]\times[k]}$ and $w_{2[N]\times[k]}$ appearing in Eq. (2.74). The ADHM constraints Eq. (2.73) are automatically satisfied since the submatrices a'_n are Hermitian. The remaining constraints, in Eq. (2.72), then form the non-trivial $U(N)$ k-instanton ADHM constraints. The constraints Eq. (2.73) can now be written in a yet more explicit form using the elements of a given in Eq. (2.74). In terms of commutators

$[A, B] \equiv AB - BA$, the ADHM constraints assume their penultimate form:

$$\bar{w}_1 w_2 + [(a'_4 - ia'_3), (a'_2 + ia'_1)] \;=\; 0, \tag{2.75}$$

$$\bar{w}_1 w_1 - \bar{w}_2 w_2 + [(a'_4 - ia'_3), (a'_4 + ia'_3)] + [(a'_2 + ia'_1), (a'_2 - ia'_1)] \;=\; 0. \tag{2.76}$$

The first constraint, Eq. (2.75), and second constraint, Eq. (2.76), respectively, are often referred to as the 'complex ADHM constraint,' and the 'real ADHM constraint.' For topological charge $k \geq 2$, however, both of these matrix equations will contain real and complex elements.

The most explicit form of the $U(N)$ k-instanton ADHM constraints is obtained by substituting the explicit forms of the matrices a'_n and w_α in Eqs. (2.75,2.76). The number of elements contained within these matrices depends on the instanton charge k and therefore the ADHM constraints differ for different values of k. The number of explicit ADHM constraints obtained from Eqs. (2.75,2.76) increases with increasing k.

We can now count the number of real independent parameters which solutions of the ADHM constraints Eqs. (2.75,2.76) shall possess. The ADHM matrix a, in the form given in Eq. (2.74) contains $4Nk + 4k^2$ real parameters. The ADHM constraints in Eq. (2.76) then place $3k^2$ real conditions on the elements of a. The auxiliary $U(k)$ symmetry, acting as global $U(k)$ gauge rotations, removes a further k^2 real parameters. The total number of real collective co-ordinates for the $U(N)$ ADHM k-instanton is then:

$$4Nk + 4k^2 - 3k^2 - k^2 = 4Nk, \tag{2.77}$$

This counting agrees with the result for the dimensions of the $U(N)$ k-instanton moduli space, $\mathfrak{M}_k$, given in Eq. (2.37). This confirms that the required number of local co-ordinates for $\mathfrak{M}_k$ are correctly given by the collective co-ordinates of the instanton solutions provided by the ADHM construction.

The number of collective co-ordinates can now be made explicit. In the $U(N)$ ADHM k-instanton construction, there are k 4-vector centre of mass co-ordinates, giving $4k$ position co-ordinates. In addition there are k real scale sizes, which shall be identified in Subsection 2.3.2. Finally there are $(4N-5)k$ gauge iso-orientations, which specify orientation of the $U(2)$ k-instanton in $U(N)$ group space. The total number of real physical collective co-ordinates is then:

$$(4N-5)k + 4k + k \;=\; 4Nk - 5k + 4k + k \;=\; 4Nk, \tag{2.78}$$

which agrees with the total number of real collective co-ordinates given in Eq. (2.77) derived from the ADHM construction and verified by a variety of other checks.

To obtain a solution with purely physical degrees of freedom, that is, only the true collective co-ordinates required to specify the position, size and iso-orientation (in group space) of the k-instanton, the global gauge rotations are removed, leaving the number of independent physical parameters as $4Nk - N^2 + 1$ for $k \geq \frac{1}{2}N$, and $4k^2 + 1$ for $k \leq \frac{1}{2}N$ [26, 33]. For the purposes of using ADHM instanton configurations in instanton calculus, however, the global gauge rotations must be included, since they appear in the k-instanton measure and are to be integrated over [47, 213]. We defer description of the instanton calculus and its applications to Chapter 6.

Symmetries of the ADHM Construction

Instanton solutions break symmetries of the classical gauge theory, and here we describe these symmetries in terms of the moduli space $\mathfrak{M}_k$ as derived from the ADHM construction. Firstly, the action of the conformal group of Yang–Mills gauge field theory on the instanton configuration can be expressed in terms of the ADHM matrices a and b. The action of the conformal transformation upon the spacetime co-ordinate x was given in Eq. (2.33). In the same quaternionic notation of Eq. (2.33), the conformal group acts upon the ADHM matrix $\Delta(x; a, b)$ as:

$$\Delta(x'; a, b) \;=\; \Delta(x; aD + bB, aC + bA)(Cx + D)^{-1}, \tag{2.79}$$

where, as before, $\{A, B, C, D\} \in \mathbb{H}$ and $AD - BC = 1$. The factor of $(Cx + D)^{-1}$ is redundant because the gauge field depends only on the matrices U and $\bar{U}$; the action of the conformal group upon a and b is then:

$$a \rightarrow aD + bB, \quad b \rightarrow aC + bA. \tag{2.80}$$

This action can also be given for the canonical forms of a and b by implementing the transformations which take a and b to its canonical form in Eq. (2.80).

Spacetime translations are contained within the conformal group and the effect of such translations upon the canonical ADHM matrices a and b can be made explicit. The

ADHM matrix $\Delta(x; a, b)$ transforms under a translation $x \to x + \epsilon$ as:

$$\Delta(x+\epsilon; a, b) \;=\; \Delta(x; a + b\epsilon, b). \tag{2.81}$$

The translation in Eq. (2.81) then implies that the submatrices a'_n of a transform according to:

$$a'_n \;\to\; a'_n + \epsilon_n 1_{[k]\times[k]}, \quad w_{\dot{\alpha}} \;\to\; w_{\dot{\alpha}}. \tag{2.82}$$

The transformations in Eq. (2.82) enable one to identify the co-ordinates at which the centre of the instanton, or more precisely, the centre of mass of the instanton, is positioned. These degrees of freedom are termed the translational collective co-ordinates, by which one can translate the position of the instanton in spacetime. In terms of the quaternionic notation, these degrees of freedom are the four real co-ordinates contained in $X_{\alpha\dot{\alpha}} = X_n \sigma^n_{\alpha\dot{\alpha}}$, where:

$$a_{\lambda j \dot{\alpha}} \;=\; -b^{\alpha}_{\lambda j} X_{\alpha\dot{\alpha}}. \tag{2.83}$$

The relation in Eq. (2.83) can also be expressed without quaternionic notation. In this case, the instanton centre of mass co-ordinates X_n are the components of the submatrix a' proportional to the unit matrix $1_{[k]\times[k]}$, which are given by:

$$(X_i)_n \;=\; -(a'_n)_{ii}, \tag{2.84}$$

in which there is no sum over the instanton number index i ($i = 1, \ldots, k$), as this index labels separate one-instanton configurations. The co-ordinates X_n do not enter into the ADHM constraints in Eq. (2.73) and this is directly related to the product which the moduli space $\mathfrak{M}_k$ can be decomposed into in Eq. (2.30).

The other gauge theory symmetries on the instanton moduli space which we shall describe are global gauge transformations. These transformations affect only objects which have the gauge group indices $u, v, \ldots$, in the ADHM construction. Hence we expect only the submatrices $w = w_{ui\dot{\alpha}}$ to transform under the action of the gauge group $U(N)$. The submatrices w are, in terms of the ADHM index decomposition, a set of $2k$ N-vectors (or $2k$ $N \times 1$ matrices), with complex elements. If $N \leq 2k$, the global gauge transformations act non-trivially upon the ADHM matrices. If $N > 2k$ there exists a subgroup whose action leaves the instanton configuration fixed. This subgroup is non-trivial and is known as the stability group of the instanton.

One can identify the stability group of the instanton by embedding the generic k-instanton configuration into a suitable subgroup of the gauge group [26]. For gauge group $SU(N)$, this subgroup is $SU(2k) \subset SU(N)$, if $N > 2k$. The embedding uses a gauge transformation $\mathfrak{U}$ which puts the $2k \times N$ submatrix w, with elements $w_{ui\dot{\alpha}}$, into upper triangular form, denoted w_{tri}, via $w = \mathfrak{U} \cdot w_{\text{tri}}$.

Using this embedding, an $SU(N)$ k-instanton solution $v_m^{\text{k-inst}}$ which has the stability group taken into account will then have the generic form given by:

$$v_m = \mathfrak{U}^\dagger \begin{pmatrix} v^{\text{k-inst}}_{m[2k]\times[2k]} & 0_{[2k]\times[N-2k]} \\ 0_{[N-2k]\times[2k]} & 0_{[N-2k]\times[N-2k]} \end{pmatrix} \mathfrak{U}, \tag{2.85}$$

where, for $N > 2k$, the element $\mathfrak{U}$ belongs to the stability group:

$$\mathfrak{U} \in \frac{SU(N)}{S\left(U(N-2k) \times U(1)\right)}. \tag{2.86}$$

Thus $\mathfrak{U}$ implements non-trivial global gauge transformations on the instanton configuration. When $N \leq 2k$, one has the usual global gauge transformations which act non-trivially on the instanton, for which the stability group would be $\mathfrak{U} \in SU(N)$. The element $\mathfrak{U}$ can be considered a 'gauge orientation' for the instanton solution, although for $N > 2k$, the group to which $\mathfrak{U}$ belongs must be quotiented as in Eq. (2.86). Also, the $U(1)$ group involved in the quotient Eq. (2.86) can also be identified with the $U(1)$ subgroup of the residual symmetry group $U(k)$ in the ADHM construction.

The $U(N)$ One-Instanton

We now use the ADHM construction to explicitly derive the most general self-dual $SU(N)$ or $U(N)$ gauge field with topological charge $k = 1$. This gauge field will be the general $U(N)$ ADHM one-instanton gauge field. To do this requires one to determine the ADHM matrix $U(x)$ in Eq. (2.42) from the ADHM construction. Following [223], we can use the ADHM index decomposition to write the ADHM matrices U and Δ as:

$$U_{\lambda v} = U_{(u+i\alpha)v} = \begin{pmatrix} V_{uv} \\ (U'_\alpha)_{iv} \end{pmatrix}, \quad \Delta_{\lambda j\dot{\alpha}} = \Delta_{(u+i\alpha)j\dot{\alpha}} = \begin{pmatrix} w_{uj\dot{\alpha}} \\ (\Delta'_{\alpha\dot{\alpha}})_{ij} \end{pmatrix}. \tag{2.87}$$

The dimensions of the various submatrices in Eq. (2.87) can be made explicit as follows:

$$U_{[N+2k]\times[N]} = \begin{pmatrix} V_{[N]\times[N]} \\ U'_{[2k]\times[N]} \end{pmatrix}, \quad \Delta_{[N+2k]\times[2k]} = \begin{pmatrix} w_{[N]\times[2k]} \\ \Delta'_{[2k]\times[2k]} \end{pmatrix}. \tag{2.88}$$

The matrix U can be determined in terms of the fundamental ADHM matrix Δ, which depends on the matrices a and b. If a and b are in canonical form, then knowledge of the matrix a derived from solving the ADHM constraints will then enable one to explicitly find Δ. From Δ, as will be shown, U, and thus the gauge field v_m, can be explicitly determined.

Substituting Eq. (2.88) into the completeness relation in Eq. (2.44), one has:

$$V_{[N]\times\underline{[N]}}\bar{V}_{\underline{[N]}\times[N]} = 1_{[N]\times[N]} - w_{[N]\times\underline{[k]}\times\underline{[2]}}f_{\underline{[k]}\times\underline{[k]}}\bar{w}_{\underline{[2]}\times\underline{[k]}\times[N]}, \tag{2.89}$$

or more succinctly, using Eq. (2.87), this can be written as:

$$V\bar{V} \;=\; V^2 \;=\; 1_{[N]\times[N]} - w_{\dot{\alpha}}f\bar{w}^{\dot{\alpha}}, \tag{2.90}$$

where we have used the Hermicity of V, which follows from the Hermicity of f. Any matrices V which solve Eq. (2.90) are related to each other by the local gauge transformation $V \to Vg_N(x)$, where $g_N(x) \in U(N)$ is an x dependent matrix. Selecting a particular V is therefore associated with fixing the local spacetime gauge of the instanton. Following [223], we choose to work in 'singular gauge,' in which V is given by one of the matrix square roots of Eq. (2.90):

$$V = (1 - w_\alpha f\bar{w}^{\dot{\alpha}})^{1/2}, \tag{2.91}$$

of which there are 2^N choices in number. In general V will be a matrix. To take the matrix square root in Eq. (2.91), one can diagonalise V^2 and take the square roots of the diagonal elements. The result will then give V in a convenient diagonal matrix form. Again using Eq. (2.44) the submatrix U' of U can then be expressed in terms of V as follows:

$$U' = -\Delta'_\alpha f\bar{w}^{\dot{\alpha}}\bar{V}^{-1}. \tag{2.92}$$

To construct the general $U(N)$ one-instanton solution, the topological charge is set to $k = 1$ and the instanton number indices i, j in the above procedure can be dropped.

The ADHM constraints Eq. (2.72) then imply that a'_n is a real 4-vector, which can be identified with the (negative) centre of the instanton $-X_n$, via Eq. (2.84) with $k = 1$:

$$a'_n \equiv -X_n \in \mathbb{R}^4. \tag{2.93}$$

With this identification of a'_n, the ADHM constraints in Eq. (2.72) can then be written as:

$$\bar{w}^{\dot{\alpha}}_u w_{u\dot{\beta}} = \rho^2 \delta^{\dot{\alpha}}_{\dot{\beta}}. \tag{2.94}$$

The real parameter ρ can be identified with the scale size for the one-instanton, as shall be described below. The ADHM constraint in Eq. (2.94) can be readily solved for w:

$$w_{u\dot{\alpha}} = \rho\, \mathfrak{U}_{[N]\times[N]} \begin{pmatrix} 1_{[2]\times[2]} \\ 0_{[N-2]\times[2]} \end{pmatrix} \tag{2.95}$$

From the embedding described above, the stability group for the $U(N)$ one-instanton is found to be $\mathfrak{U} \in SU(N)$. Using Eq. (2.47), the explicit form for the ADHM matrix $\Delta(x)$ can now be established. In this case, it reads as:

$$\Delta(x)_{[N+2]\times[2]} = \begin{pmatrix} \rho \cdot \mathfrak{U}_{[N]\times[N]} \\ 0_{[N-2]\times[2]} \\ (x - X)_{[2]\times[2]} \end{pmatrix}. \tag{2.96}$$

To proceed further, the explicit form of the arbitrary x dependent function f can be obtained from Eq. (2.68), using Δ in Eq. (2.96) above, and is the following scalar quantity:

$$f = \frac{1}{(x - X)^2 + \rho^2}, \tag{2.97}$$

Substituting Eq. (2.97) for f into Eqs. (2.91,2.92) for V and U' then gives:

$$V = 1_{[N]\times[N]} + \frac{1}{\rho^2}\left[\sqrt{\frac{(x-X)^2}{(x-X)^2+\rho^2}} - 1\right] w_{\dot{\alpha}} \bar{w}^{\dot{\alpha}}, \tag{2.98}$$

$$U' = -\frac{(x-X)_{\alpha\dot{\alpha}} \bar{w}^{\dot{\alpha}}}{|x - X|\sqrt{(x-X)^2+\rho^2}}. \tag{2.99}$$

Substituting Eqs. (2.98,2.99) into Eq. (2.87) gives the matrix $U(x)$. Then, using the explicit form of U obtained in Eq. (2.87), one arrives at the explicit form of the $U(N)$ one-instanton gauge field in singular gauge:

$$v_m = \frac{2}{g} \frac{\mathfrak{U} w_{\dot{\alpha}} (x-X)_n \bar{\sigma}^{\dot{\alpha}}_{mn\dot{\beta}} \bar{w}^{\dot{\beta}} \mathfrak{U}^\dagger}{(x-X)^2[(x-X)^2+\rho^2]}, \tag{2.100}$$

in which the submatrices $w_{\dot{\alpha}}$ and $\bar{w}^{\dot{\beta}}$ give ρ^2 as in Eq. (2.94). The $U(N)$ one-instanton in Eq. (2.100) is the generalization of the singular gauge BPST one-instanton in Eq. (2.24), from gauge group $SU(2)$ to $SU(N)$ (we recall that the ADHM construction does not distinguish between the $U(N)$ and $SU(N)$ gauge groups). As can be seen in Eq. (2.100), the $SU(N)$ one-instanton solution is given by the $SU(2)$ BPST one-instanton solution embedded in the larger gauge group $SU(2) \subset SU(N)$. The $SU(N)$ one-instanton solution can then be written as:

$$v_m = \mathfrak{U} \begin{pmatrix} v_m^{\mathrm{BPST}} & 0 \\ 0 & 0 \end{pmatrix} \mathfrak{U}^\dagger, \tag{2.101}$$

where $\mathfrak{U}$ is now interpreted as the gauge orientation (or 'iso-orientation') of the $SU(2)$ BPST one-instanton in $SU(N)$ group space. For $k = 1$ and $N > 2$, we note that element $\mathfrak{U}$ will be a member of the coset group $SU(N)/S\left(U(N-2) \times U(1)\right)$. This is because for $k = 1$ and $N > 2$, the gauge orientation cannot be the element $\mathfrak{U} \in SU(N)$. Furthermore, for $k = 1$, if $N < 2$, the gauge group is $U(1)$, giving an Abelian gauge theory in which Yang–Mills instantons do not exist. The $SU(2)$ generators for this embedding $SU(2) \subset SU(N)$ are given by:

$$T^c_{uv} = \rho^{-2} w_{u\dot{\alpha}} \tau^{c\dot{\alpha}}_{\ \dot{\beta}} \bar{w}^{\dot{\beta}}_{v}, \tag{2.102}$$

where the index $c = 1, 2, 3$ labels the three standard Pauli matrices. The generators T^c_{uv} satisfy the $SU(2)$ Lie algebra by virtue of the ADHM constraints Eq. (2.72).

The ADHM Hyper-Kähler Quotient

The ADHM construction of instantons has a modern mathematical interpretation as being a particular instance of a construction known as the hyper-Kähler quotient [67]. In the $U(N)$ ADHM construction, the moduli space $\mathfrak{M}_k$ is described in terms of the matrices $a_{\dot{\alpha}}$ upon which the ADHM constraints Eqs. (2.72,2.73) are imposed, followed by a quotient of the space of solutions of these constraints by the residual $U(k)$ symmetry group. This is an example of a hyper-Kähler quotient, which was first noticed for the ADHM construction in [67]. In the following we do not describe the hyper-Kähler quotient construction in detail, but refer the reader to the review [224], which includes discussions of this mathematical topic in relation to the instanton moduli space.

In the hyper-Kähler quotient construction, one derives a hyper-Kähler space $\mathfrak{M}$ from the quotient of another hyper-Kähler space, $\tilde{\mathfrak{M}}$. The space $\tilde{\mathfrak{M}}$ is referred to as the 'mother space' and must possess appropriate isometries. For the ADHM construction, one has $\tilde{\mathfrak{M}} = \mathbb{R}^{4k(N+k)}$, and by quotienting this space with the isometry group $U(k)$, the ADHM constraints can be derived. Specifically, the ADHM constraints Eqs. (2.72,2.73) appear as the condition for the vanishing of moment maps arising from the isometry group $U(k)$. Hence the instanton moduli space $\mathfrak{M}_k$ appears naturally as a type of hyper-Kähler quotient.

The hyper-Kähler quotient construction is useful for obtaining the geometric characteristics of $\mathfrak{M}_k$. For example, the metric on the moduli space can be directly derived from the mother space $\tilde{\mathfrak{M}}$ using the hyper-Kähler quotient construction [67].

2.3.2 $U(N)$ ADHM Multi-Instantons

In this subsection we present the first $U(N)$ ADHM multi-instanton solution. The ADHM constraints have been explicitly solved and the general form of the $U(N)$ ADHM two-instanton determined [36]. In this subsection we also consider the ADHM constraints for the $U(N)$ three-instanton.

We first describe the completely clustered, or dilute instanton gas, limit of $U(N)$ multi-instantons. This is an asymptotic limit which is an important physical property of the multi-instanton moduli space. Next the explicit general $U(N)$ ADHM two-instanton solution, first reported by the author in [36], is presented. An outline of the method used to solve the ADHM constraints for this case is given. Previously, the only known explicit general two-instanton solution of the (anti)-self-dual Yang–Mills equations was the $Sp(1)$ ADHM two-instanton [22]. This solution has seen widespread use in instanton calculations which make use of the isomorphism $Sp(1) \simeq SU(2)$. The explicit parameterization of the general $U(N)$ ADHM two-instanton given here is the first general multi-instanton configuration with unitary gauge group. In this subsection we also briefly describe the use of the ADHM construction to construct the explicit $U(N)$ two-instanton gauge field. An explicit manifestation of the dilute instanton gas limit is given for the $U(2)$ two-instanton solution. Details of the $U(2)$ residual symmetry transformations utilized in the construction of the $U(N)$ two-instanton configuration are also given.

Finally, a discussion of the $U(N)$ ADHM constraints for the case $k = 3$ is included. The ADHM constraints for $U(N)$ three-instantons possess a much greater complexity than the constraints for topological charge $k = 2$, and there are also a greater number of constraints which must be solved simultaneously. We also make some observations regarding the $U(N)$ ADHM constraints for $k > 3$.

The Dilute Instanton Gas Limit

Multi-instanton solutions are instantons which exist independently of one-instantons; in general, they do not exist as a simple sum of one-instantons. However, on the instanton moduli space $\mathfrak{M}_k$ there exist regions where such an interpretation of a $U(N)$ multi-instanton is valid. In these regions, a multi-instanton can be decomposed into a set of widely separated one-instantons. This is the completely clustered limit, or clustering limit, of the multi-instanton [224]. This limit is also known as the dilute instanton gas limit, in which a k-instanton configuration can be approximately described as a sum of k widely separated one-instantons. Other clustering limits exist for particular multi-instantons: a three-instanton would possess regions of the moduli space in which it can be described as a two-instanton plus a one-instanton as well as three one-instantons. These limits are special and depend on the multi-instanton solution being considered. However, the completely clustered limit is common to all multi-instantons.

In the ADHM construction, the completely clustered limit is the region of the instanton moduli space where the diagonal elements of the submatrices a'_n are much larger than their off-diagonal elements, at least in the conventional physical interpretation of known ADHM instanton configurations. In this limit, the centres of the k one-instantons can be identified as the co-ordinates $(X_n)_i \equiv -(a'_n)_{ii}$, $i = 1, \ldots, k$. This limit is defined up to the action of the residual $U(k)$ symmetry. One can use the $U(k)$ symmetry to set off-diagonal components of a'_n which arise due to the action of the $U(k)$ group on the diagonal matrix $\mathrm{diag}(-(X_n)_1, \ldots, (X_n)_k)$ to zero. With this use of the $U(k)$ symmetry, the result is that the submatrices a'_n are constrained as:

$$(a'_n)_{ij}\,(X_n)_i - (X_n)_j) \;=\; 0. \tag{2.103}$$

The dilute instanton gas limit can also be expressed as a condition involving only the submatrices a'_n [225]:

$$[a'_n\;,\;a'_m] \;\to\; 0, \forall\; m,n. \tag{2.104}$$

The condition Eq. (2.104) implies that there exists a $U(k)$ residual symmetry transformation which diagonalizes each of the submatrices a'_n, so that they can be written as $a'_n = \mathrm{diag}[(a'_n)_{11}, \ldots, (a'_n)_{kk}]$, for instanton number k. This assists in the identification of the instanton centre of mass co-ordinates $(X_n)_i$ stated above.
We denote the submatrices a'_n constrained as in Eqs. (2.103,2.104) as $\tilde{a}'_n$. The constraint Eq. (2.103) gives rise to a diagonal symmetry $U(1)^k$. Each $U(1)$ factor in the $U(1)^k$ symmetry can be identified with the residual $U(k) = U(1)$ symmetry of the widely separated one-instantons in the clustering limit. In this limit the quadratic term $(\tilde{a}'^{\dot{\alpha}\alpha})_{ik}(\tilde{a}'_{\alpha\dot{\beta}})_{kj}$, $k \neq i,j$ of the ADHM constraint Eq. (2.73) can be ignored. The ADHM constraints Eq. (2.73) then have off-diagonal components linear in $(\tilde{a}'_{\alpha\dot{\alpha}})_{ij}$, $i \neq j$, given by:

$$(\bar{X}^i - \bar{X}^j)^{\dot{\alpha}\alpha}(\tilde{a}'_{\alpha\dot{\beta}})_{ij} + \bar{w}^{\dot{\alpha}}_{iu} w_{uj\dot{\beta}} \;\propto\; \delta^{\dot{\alpha}}_{\dot{\beta}}, \tag{2.105}$$

which can be explicitly solved for the constrained diagonal elements $(\tilde{a}'_{\alpha\dot{\alpha}})_{ij}$. The diagonal components of the ADHM constraints Eq. (2.73) are given by:

$$\bar{w}^{\dot{\alpha}}_{iu} w_{ui\dot{\beta}} \;=\; \rho_i^2 \delta^{\dot{\alpha}}_{\dot{\beta}}, \tag{2.106}$$

in which there is no summation on the instanton number index i. The parameters $\{\rho_i\}$ are arbitrary and real. Equation (2.106) is the generalization of Eq. (2.94) for arbitrary instanton number k, and can be identified as a set of k one-instanton constraints. This implies that the k parameters $\{\rho_i\}$ can also be identified as the scale sizes of the one-instantons, with ρ_i being the scale size of the i^{th} one-instanton. The parameters $\{\rho_i\}$ can be expressed explicitly in terms of the submatrices $w_{ui\dot{\alpha}}$ as [225]:

$$\rho_i^2 \;=\; \tfrac{1}{2}\bar{w}^{\dot{\beta}}_{iu} w_{ui\dot{\beta}}. \tag{2.107}$$

Each of the one-instantons also has an interpretation as a gauge oriented $SU(2)$ one-instanton embedded in the completely clustered $SU(N)$ multi-instanton. The generators which give these embeddings are given by:

$$(T_i^c)_{uv} = \rho_i^{-2} w_{ui\dot{\alpha}} \tau_{\dot{\beta}}^{c\dot{\alpha}} \bar{w}_{iv}^{\dot{\beta}}, \tag{2.108}$$

with no sum over the instanton number index i. Equation (2.108) generalizes the set of one-instanton $SU(2)$ generators in Eq. (2.102) to the case of k-instantons. In terms of the generators $(T_i^c)_{uv}$, the completely clustered limit applies when, for each instanton number index $i \neq j$, one has:

$$(X_i - X_j)^2 \gg \rho_i \rho_j \mathrm{tr}_N(T_i^c T_j^c), \tag{2.109}$$

with no sum over the instanton number index i. This condition implies that in the completely clustered limit for $U(N)$ multi-instantons, any two of the one-instantons contained within the multi-instanton configuration must be separated at a distance much greater than the product of the scale sizes multipled by the trace over the product of generators $(T_i^c T_j^c)$. This product of the generators quantifies the extent to which the individual $SU(2) \subset SU(N)$ embeddings of the given pair of one-instantons labelled i and j overlap one another.

The $U(N)$ ADHM Two-Instanton

With the aim of determining the most general solution of the $U(N)$ ADHM constraints for $k = 2$, which will have $4Nk = 8N$ real independent parameters, we firstly need to find a solution with $(8N + 4)$ real parameters. This then allows one to rotate out the $U(k) = U(2)$ residual symmetry, effectively eliminating four real parameters in the $(8N + 4)$-parameter solution. We adopt the notation and conventions of the previous paragraph and Subsection 2.3.1. For later convenience, we define the following quantities in terms of the ADHM data present in the matrix a, for $k = 2$:

$$\bar{w}_1 w_2 = \sum_{n=1}^{N} \begin{pmatrix} \bar{w}_{n1,1} w_{2,n1} & \bar{w}_{n1,1} w_{2,n2} \\ \bar{w}_{n2,1} w_{2,n1} & \bar{w}_{n2,1} w_{2,n2} \end{pmatrix} \equiv \begin{pmatrix} U_x & U_y \\ U_z & U_t \end{pmatrix}, \tag{2.110}$$

$$\bar{w}_1 w_1 - \bar{w}_2 w_2 = \sum_{n=1}^{N} \begin{pmatrix} |w_{1,n1}|^2 - |w_{2,n1}|^2 & \bar{w}_{1,n1} w_{1,n2} - \bar{w}_{2,n1} w_{2,n2} \\ w_{n1,1}\bar{w}_{n2,1} - w_{n1,2}\bar{w}_{n2,2} & |w_{1,n2}|^2 - |w_{2,n2}|^2 \end{pmatrix} \equiv \begin{pmatrix} U_1 & U_2 \\ \bar{U}_2 & U_4 \end{pmatrix}. \tag{2.111}$$

Note that the sums in Eqs. (2.110,2.111) run from 1 to N; although the number N in these sums is related to the rank of the gauge group $U(N)$, the construction of ADHM instanton configurations breaks down for $N = 1$, as is to be expected given that (commutative) ADHM instantons are a phenomenon of non-Abelian gauge theories. We now make a change of variables which affects only the diagonal elements of the complex matrices $(a'_4 + ia'_3)$ and $(a'_2 + ia'_1)$, such that:

$$(a'_4 + ia'_3) = \begin{pmatrix} a + \frac{1}{2}X_1 & b \\ c & a - \frac{1}{2}X_1 \end{pmatrix}, \tag{2.112}$$

$$(a'_2 + ia'_1) = \begin{pmatrix} \alpha + \frac{1}{2}X_2 & \beta \\ \gamma & \alpha - \frac{1}{2}X_2 \end{pmatrix}, \tag{2.113}$$

where $\{X_1, X_2, a, b, c, \alpha, \beta, \gamma\} \in \mathbb{C}$. This is to make the physical interpretation of the $U(N)$ $k = 2$ ADHM instanton configurations more transparent, and to simplify calculations involving the elements of $(a'_4 + ia'_3)$ and $(a'_2 + ia'_1)$.

With the explicit form of the submatrices a'_n given above in Eq. (2.113), and the combinations involving the other submatrices $w_{\dot\alpha}$, we can now write out the algebraic ADHM constraints Eq. (2.75,2.76) for $k = 2$. These constraints are explicitly given by:

$$U_x + \bar{c}\gamma - \beta\bar{b} = 0, \tag{2.114}$$

$$U_y + \beta\bar{X}_1 - \bar{c}X_2 = 0, \tag{2.115}$$

$$U_z + \bar{b}X_2 - \gamma\bar{X}_1 = 0, \tag{2.116}$$

$$U_x + U_t = 0, \tag{2.117}$$

$$U_2 + b\bar{X}_1 - \bar{c}X_1 + \bar{\gamma}X_2 - \beta\bar{X}_2 = 0, \tag{2.118}$$

$$U_1 + |c|^2 - |b|^2 + |\beta|^2 - |\gamma|^2 = 0, \tag{2.119}$$

$$U_1 + U_4 = 0. \tag{2.120}$$

Equations (2.114–2.117) are the four distinct equations which originate from the first part of the ADHM constraints, Eq. (2.75). The remaining constraints, Eqs. (2.117–2.120) are the three distinct equations which originate from the second part of the ADHM constraints, Eq. (2.76). We note that Eq. (2.117) and Eq. (2.120) may be obtained by taking the trace over the instanton number indices i, j of the ADHM constraints Eq. (2.75) and Eq. (2.76), respectively. Our aim for the solution of the constraints is to eliminate the off-diagonal elements $\{b, c\}$ and $\{\beta, \gamma\}$ from the submatrices $(a'_4 + ia'_3)$ and $(a'_2 + ia'_1)$, respectively. Then the remaining diagonal elements, which will be proportional to b, and hence x, in quaternionic form, will contain the translational co-ordinates $\{a, d, \alpha, \delta\}$ and the instanton centre of mass co-ordinates X_1, X_2. The remaining collective co-ordinates will then be functions of the elements of the submatrices w_1 and w_2.

The constraints in Eqs. (2.114–2.120) can be considered as follows. Two of the constraints are independent of the variables of the submatrices a'_n, Eqs. (2.117,2.120). Therefore none of the off-diagonal elements $\{b, c, \beta, \gamma\}$ can be eliminated from these equations. Furthermore, Eqs. (2.118) and Eq. (2.120) are real equations. Since the variables $\{b, c, \beta, \gamma\}$ are complex, we shall not be able to completely eliminate these variables using Eqs. (2.118,2.120). This leaves the four complex equations Eqs. (2.114,2.115,2.116) and Eq. (2.118) with which to completely eliminate the set of variables $\{b, c, \beta, \gamma\}$.

We now set about solving the system of simultaneous equations Eqs. (2.114-2.120). After some trial and error involving the sequence of variables to be eliminated, the $k = 2$ $U(N)$ ADHM constraints were solved by exploiting the linearity present in them. Thus the $k = 2$ constraints were solved using only linear algebra. This is precisely the primary utility of using the ADHM construction to obtain solutions of the self-dual Yang–Mills field equations. However, the actual method of solution is not obvious from the form of the constraints. Firstly, we write the variables $\bar{b}$ and $\bar{c}$, of the submatrix $(a'_4 + ia'_3)$, in terms of the variables β and γ, of $(a'_2 + ia'_1)$, using Eqs. (2.115,2.116):

$$\bar{b} = \frac{1}{X_2}[\gamma \bar{X}_1 - U_z], \tag{2.121}$$

$$\bar{c} = \frac{1}{X_2}[\beta \bar{X}_1 + U_y]. \tag{2.122}$$

This choice does not involve the bilinear constraint Eq. (2.114) nor the real quadratic constraint Eq. (2.119). We then substitute for b and c in Eq. (2.114) and Eq. (2.118) using Eqs. (2.121,2.122). A highly fortuitous cancellation in Eq. (2.114) leads to no bilinear terms involving β and γ. The result is two complex equations linear in β and γ:

$$X_2 U_x + \gamma U_y - \beta U_z = 0, \tag{2.123}$$

$$|X_2|^2 U_2 + |X|^2 \bar{X}_2 \beta - |X|^2 X_2 \bar{\gamma} - \bar{X}_1 X_2 \bar{U}_z - X_1 \bar{X}_2 U_y = 0, \tag{2.124}$$

where we have defined the quantity $|X|^2$:

$$|X|^2 \equiv |X_1|^2 - |X_2|^2. \tag{2.125}$$

We now use Eqs. (2.123,2.124) to eliminate β and γ in terms of X_1, X_2 and the other variables dependent upon $w_{\dot{\alpha},iu}$. By writing β in Eq. (2.123) in terms of γ, one can eliminate β from Eq. (2.124). One then obtains a single equation in terms of γ and its complex conjugate, $\bar{\gamma}$:

$$|X_2|^2U_2 + |X|^2\bar{X}_2\frac{U_y}{U_z}\gamma - |X|^2X_2\bar{\gamma} + |X|^2|X_2|^2\frac{U_x}{U_z} - \bar{X}_1X_2\bar{U}_z - X_1\bar{X}_2U_y = 0. \tag{2.126}$$

Since Eq. (2.126) is linear in both γ and $\bar{\gamma}$, one can decompose Eq. (2.126) into real and imaginary parts. After this is done, we obtain two real simultaneous equations linear in $\mathrm{Re}(\gamma)$ and $\mathrm{Im}(\gamma)$. Solving these simultaneously then enables one to explicitly solve for γ in terms of the quantities X_1, X_2 and $w_{\dot{\alpha},iu}$ only. Using the previous relations Eqs. (2.115,2.116,2.123), which express the other variables $\{b, c, \beta\}$ in terms of γ, one can then write all of the off-diagonal elements $\{b, c, \beta, \gamma\}$ of a'_n in terms of X_1, X_2 and $w_{\dot{\alpha},iu}$ only.

Combining all of these results leads to the solution of the $U(N)$ $k = 2$ ADHM constraints with $(8N + 4)$ real parameters, which is given by:

$$(a'_4 + ia'_3) = \begin{pmatrix} a + \frac{1}{2}X_1 & \dfrac{X_1[\bar{P}u - P]}{|X|^2|X_2|^2(|u|^2-1)} - \dfrac{\bar{U}_z}{X_1} \\ \dfrac{\bar{U}_y}{X_1} - \dfrac{X_1\bar{U}_x}{U_z} - \dfrac{X_1\bar{u}[\bar{P}u - P]}{|X|^2|X_2|^2(|u|^2-1)} & a - \frac{1}{2}X_1 \end{pmatrix}, \tag{2.127}$$

$$(a'_2 + ia'_1) = \begin{pmatrix} \alpha + \frac{1}{2}X_2 & -\dfrac{X_2\bar{U}_x}{U_z} - \dfrac{X_2u[P\bar{u} - \bar{P}]}{|X|^2|X_2|^2(|u|^2-1)} \\ \dfrac{X_2[P\bar{u} - \bar{P}]}{|X|^2|X_2|^2(|u|^2-1)} & \alpha - \frac{1}{2}X_2 \end{pmatrix}, \tag{2.128}$$

in which we have defined the following quantities:

$$P \equiv \bar{X}_1X_2\bar{U}_z + X_1\bar{X}_2U_y - |X_2|^2U_2 - |X|^2|X_2|^2\frac{U_x}{U_y}, \tag{2.129}$$

$$u \equiv \frac{U_y}{U_z}, \tag{2.130}$$

The solution also includes two conditions which also arise from the ADHM constraints (Eqs. (2.75,2.76)), which when taken together with Eqs. (2.127,2.128) constitute the general solution of the $U(N)$ $k=2$ ADHM constraints:

$$U_x = -U_t, \tag{2.131}$$

$$U_1 = -U_4 = |b|^2 - |c|^2 + |\gamma|^2 - |\beta|^2. \tag{2.132}$$

Equations (2.131,2.132) solve the a'_n-independent constraints in Eqs. (2.117,2.120), and we have combined Eq. (2.119) with Eq. (2.120) in Eq. (2.132) for simplicity. We are careful to note that Eq. (2.132) may appear to contain dependence on the same variables in $w_{\dot{\alpha}}$ which the quantities U_x, U_y, U_z and U_2 also depend upon. This may be the case for the $(8N+4)$-parameter solution, but once the $U(k) = U(2)$ residual symmetry of the solution has been fixed, and the number of parameters reduced to $8N$, any such interdependence, which is potentially an obstacle to the general solution of the $k = 2$ constraints and shall be exceedingly complicated, will be removed. We now address the residual $U(2)$ symmetry of the $(8N+4)$-parameter solution Eqs. (2.127 – 2.132).

The $U(k)$ symmetry of Eq. (2.69) acts as follows on the submatrices of a:

$$\begin{pmatrix} w_1 & w_2 \end{pmatrix} \rightarrow \begin{pmatrix} w_1\Omega & w_2\Omega \end{pmatrix}, \tag{2.133}$$

$$\begin{pmatrix} (a'_4 + ia'_3) & (a'_2 + ia'_1) \\ -(a'_2 - ia'_1) & (a'_4 - ia'_3) \end{pmatrix} \rightarrow \Omega^\dagger \begin{pmatrix} (a'_4 + ia'_3) & (a'_2 + ia'_1) \\ -(a'_2 - ia'_1) & (a'_4 + ia'_3) \end{pmatrix} \Omega, \tag{2.134}$$

where $\Omega \in U(2)$ for topological charge $k = 2$. In a separate paragraph below we provide the details of a particular transformation Ω which can be used to set $U_x = 0$ and $w_{1,11} = 0$. Indeed any other element of w_1 or w_2 could be set to zero instead of $w_{1,11}$. We adopt this usage of the $U(2)$ auxiliary symmetry hereon.

We note that any other solutions of the ADHM constraints for gauge group $U(N)$ and $k = 2$ possessing $(8N+4)$ real parameters will be equivalent to the above solution in Eqs. (2.127–2.132) upon acting on it with the auxiliary $U(2)$ symmetry, since the instanton moduli space for $k = 2$ is connected [10]. Indeed, one can choose to eliminate the same variables β and γ in a different order, or eliminate the variables $\{b, c\}$ in two different orders, and obtain apparently different solutions of the $k = 2$ $U(N)$ ADHM constraints. However, the four solutions so obtained are all equivalent, as has been verified using a standard symbolic manipulation program, upon substituting each solution in turn into the $k = 2$ $U(N)$ ADHM constraints. This is to be expected as each solution is related to the other solutions via $U(2)$ transformations in Eqs. (2.133,2.134) [10].

Using the $U(2)$ transformation Ω, given below, permits one to construct the $U(N)$ $k=2$ ADHM instanton which has a definite physical interpretation. The form of the $8N$-parameter solution, using the $(8N+4)$-parameter solution Eqs. (2.127–2.132) is then:

$$(a_4' + ia_3') = \begin{pmatrix} a + \frac{1}{2}X_1 & \frac{X_1[\bar{R}u - R]}{|X|^2|X_2|^2(|u|^2-1)} - \frac{\bar{U}_z}{\bar{X}_1} \\ \frac{\bar{U}_y}{\bar{X}_1} - \frac{X_1\bar{u}[\bar{R}u - R]}{|X|^2|X_2|^2(|u|^2-1)} & a - \frac{1}{2}X_1 \end{pmatrix}, \quad (2.135)$$

$$(a_2' + ia_1') = \begin{pmatrix} \alpha + \frac{1}{2}X_2 & -\frac{X_2u[R\bar{u} - \bar{R}]}{|X|^2|X_2|^2(|u|^2-1)} \\ \frac{X_2[R\bar{u} - \bar{R}]}{|X|^2|X_2|^2(|u|^2-1)} & \alpha - \frac{1}{2}X_2 \end{pmatrix}. \quad (2.136)$$

The function R is defined as:

$$R \equiv \bar{X}_1 X_2 \bar{U}_z + X_1 \bar{X}_2 U_y - |X_2|^2 U_2, \quad (2.137)$$

and the conditions Eq. (2.131,2.132), with the modifications $U_x = 0$ and $w_{1,11} = 0$, complete the specification of the $8N$-parameter $U(N)$ $k=2$ ADHM instanton solution. For the case $N=2$, the explicit two-instanton configuration can assume a particularly simple form. Continuing with our choice of residual $U(2)$ symmetry used as $U_x = 0$ and $w_{1,11} = 0$, for the $U(2)$ two-instanton one has:

$$U_x = \bar{w}_{21,1} w_{2,21} = 0, \quad (2.138)$$

$$U_y = \bar{w}_{21,1} w_{2,22}, \quad (2.139)$$

$$U_z = \bar{w}_{12,1} w_{2,11} + \bar{w}_{22,1} w_{2,21}, \quad (2.140)$$

$$U_t = \bar{w}_{12,1} w_{2,12} + \bar{w}_{22,1} w_{2,22}, \quad (2.141)$$

One of the possible solutions of Eq. (2.139) is $w_{1,21} = 0$, and we adopt this value of $w_{1,21}$ for the $U(2)$ two-instanton. Using $w_{1,21} = 0$, one then has $U_y = 0$, and U_z and U_t remain unmodified, leaving:

$$U_2 = -(\bar{w}_{11,2} w_{2,12} + \bar{w}_{21,2} w_{2,22}), \quad (2.142)$$

$$U_1 = -|w_{2,11}|^2 - |w_{2,21}|^2. \quad (2.143)$$

With these choices, any off-diagonal elements proportional to U_y then vanish, and the matrices $(a_4' + ia_3')$ and $(a_2' + ia_1')$ in Eqs. (2.135,2.136) for $N = 2$ simplify to:

$$(a_4' + ia_3') = \begin{pmatrix} a + \frac{1}{2}X_1 & \frac{1}{|X|^2|X_2|^2}[X_2|X_1|^2\bar{U}_z - X_1|X_2|^2U_2] \\ 0 & a - \frac{1}{2}X_1 \end{pmatrix}, \quad (2.144)$$

$$(a_2' + ia_1') = \begin{pmatrix} \alpha + \frac{1}{2}X_2 & 0 \\ \frac{1}{|X|^2|X_2|^2}[X_1|X_1|^2U_z - X_2|X_2|^2\bar{U}_2] & \alpha - \frac{1}{2}X_2 \end{pmatrix}, \quad (2.145)$$

Using Eqs. (2.139–2.143) and the choice $w_{1,21} = 0$, one can choose to eliminate $w_{1,12}$ via the relation $U_x = -U_t = 0$ in (Eq. (2.131)). This then implies that U_z is proportional to $\bar{w}_{22,1}$. The modulus of $w_{1,22}$ can thus be eliminated via $U_1 = -U_4$ in Eq. (2.132). The remaining constraint, the second equality in Eq. (2.132), then enables one to eliminate the imaginary part of $w_{1,22}$ through a quadratic relation in this quantity. A similar procedure has been performed for $N = 3$, in which the constraint Eq. (2.132) becomes more involved, but other choices of elements within w_1 and w_2 to eliminate can be made in order to simplify this. The number of independent real free parameters remaining in the solution is then sixteen: eight from $\{w_{2,11}, w_{2,12}, w_{2,21}, w_{2,22}\}$ and eight from $\{a, \alpha, X_1, X_2\}$), which agrees with the general result of $8N = 16$ real parameters from the parameter counting in Eq. (2.77). Hence, upon fixing the residual $U(2)$ symmetry, the above ADHM data for the $U(2)$ two-instanton configuration represents the unique sixteen parameter solution of the ADHM constraints for the gauge group $U(2)$ and topological charge $k = 2$.

We note that physical quantities constructed from the $SU(2)$ two-instanton configuration, which can be obtained from the $U(2)$ two-instanton configuration given above, will be equivalent to those constructed from the $Sp(1)$ two-instanton [22] due to the isomorphism $SU(2) \simeq Sp(1)$.

In the case of gauge group $U(N)$, with $N > 1$, the following identification of physical parameters in the solution can be made. The instanton centre of mass co-ordinates (translational co-ordinates) are given by a and α, which are proportional to x_n and can thus be set to zero. The relative instanton positions can then be taken to be X_1 and X_2. The scale sizes can be expressed using the definition given for the $U(N)$ k-instanton scale sizes in Eq. (2.107), first given in [225], as:

$$\rho_1^2 = \tfrac{1}{2}U_4 - \tfrac{1}{2}\sum_{n=1}^{N}|w_{2,n2}|^2 = \tfrac{1}{2}\sum_{n=1}^{N}(|w_{1,n1}|^2 - |w_{2,n1}|^2), \tag{2.146}$$

$$\rho_2^2 = \tfrac{1}{2}U_1 - \tfrac{1}{2}\sum_{n=1}^{N}|w_{2,n1}|^2 = \tfrac{1}{2}\sum_{n=1}^{N}(|w_{1,n2}|^2 - |w_{2,n2}|^2). \tag{2.147}$$

The global gauge orientations $\mathfrak{U}$, which will include iso-orientations for any chosen N, are given by the remaining parameters contained within the submatrices w_1 and w_2. These parameters serve to rotate the two-instanton solution in the group space of $U(N)$; through these submatrices any $U(N)$ two-instanton can be specified, and no embedding is necessary at this stage.

Thus, for the $U(2)$ solution given above, the relative instanton separations are $\{X_1, X_2\}$, the instanton centre of mass positions are $\{a, \alpha\}$, and the two scale sizes are ρ_1 and ρ_2, as defined in Eqs. (2.146,2.147). The six $U(2)$ iso-orientations are contained in the remaining elements $\{w_{2,11}, w_{2,12}, w_{2,21}, w_{2,22}\}$ taken together with the conditions which relate them.

We can now make a count of the parameters appearing in the $U(N)$ two-instanton solution. The instanton translational co-ordinates and relative separations, $\{a, \alpha, X_1, X_2\}$, give eight real parameters. There are two scale sizes, $\{\rho_1, \rho_2\}$, given by Eqs. (2.146,2.147), which are two real parameters. Also there are $(4N-5)k = (8N-10)$ real iso-orientations. Summing these gives $(8N - 10 + 8 + 2) = 8N$ real parameters, as required by the parameter counting in Eq. (2.77).

This solution must also exhibit the correct decomposition into two constituent one-instanton configurations in the completely clustered limit, which is a physically required property. This asymptotic limit can most simply be achieved by taking the relative instanton positions to infinity, $|X_i| \to \infty$, $i = 1, 2$; that is, the separation of the two coupled one-instantons approximately comprising the two-instanton is taken to be infinite in extent. In this way, the description of the two-instanton is approximated as a non-interacting dilute gas of two one-instantons ('single instantons'). The result is that the matrix a can be explicitly decomposed for $N = 2$ as:

$$\lim_{X_1,X_2\to\infty} a_{[6]\times[4]} \to \begin{pmatrix} w_{1,11} & 0 & w_{2,11} & 0 \\ w_{1,21} & 0 & w_{2,21} & 0 \\ a+\frac{1}{2}X_1 & 0 & \alpha+\frac{1}{2}X_2 & 0 \\ 0 & 0 & 0 & 0 \\ -\bar{\alpha}-\frac{1}{2}\bar{X}_2 & 0 & \bar{a}+\frac{1}{2}\bar{X}_1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & w_{1,12} & 0 & w_{2,12} \\ 0 & w_{1,22} & 0 & w_{2,22} \\ 0 & 0 & 0 & 0 \\ 0 & a-\frac{1}{2}X_1 & 0 & \alpha-\frac{1}{2}X_2 \\ 0 & 0 & 0 & 0 \\ 0 & -\bar{\alpha}+\frac{1}{2}\bar{X}_2 & 0 & \bar{a}-\frac{1}{2}\bar{X}_1 \end{pmatrix}, \tag{2.148}$$

where we have restored $w_{1,11}$ and $w_{1,21}$ for clarity. Other alternative choices of $U(2)$ transformation can also be used.

The one-instantons are centered at (X_1, X_2) and at $(-X_1, -X_2)$, respectively, and have scale sizes ρ_1 and ρ_2. This decomposition in the dilute instanton gas limit can be extended to $U(N)$, in which case the submatrices a'_n will decompose in the same manner as in Eq. (2.148) and the submatrices w_1 and w_2 will decompose in a similar way. The $U(2)$ two-instanton may perhaps assist in uncovering a 'dictionary' relating it to the $Sp(1)$ ADHM formalism [223], thus connecting the collective co-ordinates which describe these instantons.

Construction of the $U(N)$ ADHM Two-Instanton Gauge Field

To construct the $U(N)$ ADHM two-instanton gauge field, one first determines the x-dependent Hermitian matrix $f_{[k]\times[k]}$. For $k = 2$, we use Eqs. (2.39,2.18) for constructing the matrix $\Delta = a + bx$. Then the factorization condition Eq. (2.48), expressed in terms of the elements belonging to the matrices $(a'_4 + ia'_3)$, $(a'_2 + ia'_1)$, given by Eqs. (2.135,2.136)), and w_1, w_2, is:

$$\bar{\Delta}\Delta = f^{-1}1_{[2]\times[2]} = \begin{pmatrix} f_1^{-1} & 0_{[2]\times[2]} \\ 0_{[2]\times[2]} & f_2^{-1} \end{pmatrix}, \tag{2.149}$$

where the inverse matrices f_1^{-1} and f_2^{-1} are given by:

$$f_1^{-1} = \sum_{n=1}^{N} \begin{pmatrix} |w_{1,n1}|^2 + |A_1|^2 + |B_1|^2 + |c|^2 + |\beta|^2 & \bar{w}_{n1,1} w_{1,n2} + \bar{A}_1 b + A_2 \bar{c} + B_1 \bar{\gamma} + \bar{B}_2 \beta \\ w_{1,n1} \bar{w}_{n2,1} + A_1 \bar{b} + \bar{A}_2 c + \bar{B}_1 \gamma + B_2 \bar{\beta} & |w_{1,n2}|^2 + |A_2|^2 + |B_2|^2 + |b|^2 + |\gamma|^2 \end{pmatrix},$$

$$f_2^{-1} = \sum_{n=1}^{N} \begin{pmatrix} |w_{2,n1}|^2 + |A_1|^2 + |B_2|^2 + |b|^2 + |\gamma|^2 & \bar{w}_{n1,2} w_{2,n2} + A_1 \bar{c} + \bar{A}_2 b + \bar{B}_1 \beta + B_2 \bar{\gamma} \\ w_{2,n1} \bar{w}_{n2,2} + \bar{A}_1 c + A_2 \bar{b} + B_1 \bar{\beta} + \bar{B}_2 \gamma & |w_{2,n2}|^2 + |A_2|^2 + |B_2|^2 + |c|^2 + |\beta|^2 \end{pmatrix},$$

in which we have defined the quantities:

$$A_1 \equiv a + \tfrac{1}{2} X_1 + z_1, \quad A_2 \equiv a - \tfrac{1}{2} X_1 + z_1, \quad B_1 \equiv \alpha + \tfrac{1}{2} X_2 + z_2, \quad B_2 \equiv \alpha - \tfrac{1}{2} x_1 + z_2. \quad (2.150)$$

We note that the product $f^{-1} 1_{[2]\times[2]}$ in Eq. (2.149) is a possible source of some ambiguity. The correct form of this product becomes clear when compared with the result of calculating $\bar{\Delta}\Delta$.

One can choose to invert either f_1^{-1} or f_2^{-1}, with either choice being valid. This is because the matrices f_1^{-1} and f_2^{-1} arising from Eq. (2.149) are related by the ADHM constraints. This can be seen explicitly since the equality $f_1^{-1} = f_2^{-1}$ implied by Eq. (2.149) reproduces two of the original $U(N)$ $k = 2$ ADHM constraints. Upon inverting either f_1^{-1} or f_2^{-1}, it remains to determine V and U' using the selected form of f. The matrix V in Eq. (2.91) is manifestly Hermitian due to the Hermiticity of f. From Eq. (2.90) the matrix V^2 can be calculated, yielding an $N \times N$ matrix with entries dependent on the elements of f and $\{w_1, w_2\}$. In order to determine V, one can take the square root of the matrix V^2 by first diagonalising V^2 and then taking the square root of each element in the resulting diagonal matrix. We denote the generic diagonalised matrix V as:

$$V = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}, \quad (2.151)$$

where the elements $\{\lambda_v\}$ are the square roots of the N characteristic values of the matrix V^2. In this method for performing the matrix square root in Eq. (2.91), the characteristic equation for V^2 will in general be a polynomial of degree N. Thus this method is restricted to $N \leq 5$ if the characteristic equation is to be solved by radicals. An explicit similarity transformation for diagonalising V^2 would enable this difficulty to be circumvented, but this is not a systematic solution. Alternatively, for $N \geq 5$ one could embed the above $U(N)$ $k = 2$ solution with $N = 4$ and use appropriate gauge orientations, given by the coset element $\mathfrak{U}$ for $N > 2k$ in Eq. (2.86). This embedding can then be used to specify the explicit $U(N)$ two-instanton solution with $N \geq 5$, from which one can determine V. This is perhaps the only feasible way in which to construct $U(N)$ two-instanton gauge field configurations with $N \geq 5$.

Given V as in Eq. (2.151), the matrix U' can be determined using Eq. (2.92). Equation (2.92) has the following explicit form in terms of matrix muliplication for general k:

$$U'_{[2k]\times[N]} = -\Delta'_{[2k]\times\underline{[k]}\times\underline{[2]}} f_{\underline{[k]}\times\underline{[k]}} \bar{w}_{\underline{[2]}\times\underline{[k]}\times\underline{[N]}} \bar{V}^{-1}_{\underline{[N]}\times[N]}. \tag{2.152}$$

For $k = 2$, given the form of V in Eq. (2.151), for generic N, the matrix U' derived from Eq. (2.152) reads as:

$$U'_{[4]\times[N]} = -\begin{pmatrix} U'_{11} & U'_{12} & \cdots & U'_{1N} \\ U'_{21} & U'_{22} & \cdots & U'_{2N} \\ U'_{31} & U'_{32} & \cdots & U'_{3N} \\ U'_{41} & U'_{42} & \cdots & U'_{4N} \end{pmatrix}, \tag{2.153}$$

where we have utilized the definitions in Eq. (2.150). The elements of U' in Eq. (2.153) can be now be written in terms of the elements of f, denoted by f_{ij}:

$$\begin{aligned}
U'_{1v} &= \frac{1}{\lambda_v}[A_1(\bar{w}_{v1,1}f_{11} + \bar{w}_{v2,1}f_{12}) + b(\bar{w}_{v1,1}f_{21} + \bar{w}_{v2,1}f_{22}) + \bar{w}_{v1,2}B_1 + \bar{w}_{v2,2}\beta], \\
U'_{2v} &= \frac{1}{\lambda_v}[c(\bar{w}_{v1,1}f_{11} + \bar{w}_{v2,1}f_{12}) + A_2(\bar{w}_{v1,1}f_{21} + \bar{w}_{v2,1}f_{22}) + \bar{w}_{v1,2}\gamma + \bar{w}_{v2,2}B_2], \\
U'_{3v} &= \frac{1}{\lambda_v}[-\bar{B}_1(\bar{w}_{v1,1}f_{11} + \bar{w}_{v2,1}f_{12}) - \bar{\gamma}(\bar{w}_{v1,1}f_{21} + \bar{w}_{v2,1}f_{22}) + \bar{w}_{v1,2}\bar{A}_1 + \bar{w}_{v2,2}\bar{c}], \\
U'_{4v} &= \frac{1}{\lambda_v}[-\bar{\beta}(\bar{w}_{v1,1}f_{11} + \bar{w}_{v2,1}f_{12}) - \bar{B}_2(\bar{w}_{v1,1}f_{21} + \bar{w}_{v2,1}f_{22}) + \bar{w}_{v1,2}\bar{b} + \bar{w}_{v2,2}\bar{A}_2].
\end{aligned}$$

The ADHM matrix U for $U(N)$ and $k = 2$ is then given by:

$$U_{[N+4]\times[N]} = \begin{pmatrix} V_{[N]\times[N]} \\ U'_{[4]\times[N]} \end{pmatrix}, \tag{2.154}$$

and the corresponding instanton gauge field configuration v_m follows from substituting U into Eq. (2.51).

The $U(N)$ ADHM two-instanton configuration presented here could conceivably be used in instanton calculus in a number of applications. In particular, testing the proposed exact solutions of $\mathcal{N}=2$ supersymmetric $SU(N)$ Yang–Mills gauge field theories at the two-instanton level via the supersymmetric multi-instanton calculus comprehensively developed in [213, 214, 217]. These proposed exact results, the instanton calculus and the instanton tests of the exact results will be described in Chapter 6. Matching between the proposed exact results and instanton calculations, obtained at the one-instanton level, will also be described in Chapter 6. The explicit form of the general $U(N)$ two-instanton is conceivably the first step towards extending the precise matching between the exact results and instanton predictions in $\mathcal{N}=2$ supersymmetric $SU(N)$ gauge theories to the two-instanton level. We shall describe this possible extension in more detail in Chapter 6.

Use of the $U(2)$ Residual Symmetry for the $U(N)$ ADHM Two-Instanton

Here we give the details of the $U(2)$ transformation which can be used to set $U_x=0$ and $w_{1,11}=0$ within the ADHM data for the $U(N)$ $k=2$ ADHM instanton.
Using the isomorphism $U(N)\simeq U(N-1)\times SU(N)$, we can take $\Omega\in U(2)$ to be the product of a $U(1)$ transformation and an $SU(2)$ transformation. Then the $U(1)$ factor of Ω acts trivially on the submatrices a'_n, as is evident from Eq. (2.134).
However, the $U(1)$ factor of Ω acts non-trivally upon the submatrices w_α. Writing Ω as $\Omega\simeq\Upsilon\times\Xi$, the following $U(1)$ and $SU(2)$ elements, Υ and Ξ, respectively, can be chosen in order to set $U_x=0$ and $w_{1,11}=0$:

$$\begin{aligned}\Upsilon &= e^{i\chi}1_{[2]\times[2]}\ \in U(1),\\ \Xi &= \begin{pmatrix}\rho e^{i\theta} & \sqrt{1-\rho^2}e^{i\phi}\\ -\sqrt{1-\rho^2}e^{-i\phi} & \rho e^{-i\theta}\end{pmatrix}\in SU(2),\end{aligned}$$

where we have used the quantities:

$$e^{i\chi} = \mathrm{Im}(w_{1,11}), \tag{2.155}$$

$$\begin{aligned} e^{i\theta} &= \sqrt{Q} + i\sqrt{1-Q}, && (2.156)\\ Q &\equiv \left[\frac{\mathrm{Im}(U_z)}{\mathrm{Im}(U_y)} + 1\right] \cdot \left[\frac{\rho^2 |w_{1,11}|^2}{(\mathrm{Re}(w_{1,12}))^2(1-\rho^2)} + \frac{2\,\mathrm{Im}(U_y)}{\mathrm{Im}(U_z)}\right]^{-1}, && (2.157)\end{aligned}$$

$$e^{-2i\phi} = -\frac{\mathrm{Im}(U_z)}{\mathrm{Im}(U_y)} e^{2i\theta}, \tag{2.158}$$

$$\rho^2 = \tfrac{1}{2} \pm \sqrt{1 - \frac{4\,\mathrm{Im}(U_y)\mathrm{Im}(U_z)|U_x|}{\left[4\,\mathrm{Im}(U_y)\mathrm{Im}(U_z)|U_x| - (\mathrm{Re}(U_y)\mathrm{Im}(U_z) - \mathrm{Im}(U_y)\mathrm{Re}(U_z))^2\right]}}. \tag{2.159}$$

We note that the $SU(2)$ part of the $U(2)$ symmetry would also enable one to set $\bar{U}_y = U_z$, but this choice would have rendered the general exact solution (Eqs. (2.127–2.132)) of the $k = 2$ $U(N)$ ADHM constraints singular, since if $\bar{U}_y = U_z$ then $(|w|^2 - 1)^{-1} \to \infty$. To obtain a physically meaningful solution we are thus induced to choose $U_x = 0$. Other $U(2)$ transformations can be implemented to act upon the $(8N+4)$-parameter $U(N)$ two-instanton solution given in Eqs. (2.127–2.132).

The considerations which led to the choice of $U(2)$ transformation specified by Eq. (2.155) and Eq. (2.156) involve the matrix $\bar{w}_1 w_2$ given in Eq. (2.110). We first write this matrix as:

$$\mathcal{U} \equiv \bar{w}_1 w_2 \equiv \begin{pmatrix} U_x & U_y \\ U_z & -U_x \end{pmatrix}, \tag{2.160}$$

where we have used Eq. (2.117) to write $U_t = -U_x$. We now consider the Hermitian matrix $\mathcal{U}\bar{\mathcal{U}}$. Under a $U(2)$ residual symmetry transformation, which acts upon the sub-matrices $w_{\dot{\alpha}}$ as in Eq. (2.71), $\mathcal{U}\bar{\mathcal{U}}$ transforms as:

$$\mathcal{U}\bar{\mathcal{U}} \to \Omega\mathcal{U}\bar{\mathcal{U}}\Omega^\dagger, \tag{2.161}$$

where $\mathcal{U}\bar{\mathcal{U}}$ is explicitly given by the matrix:

$$\mathcal{U}\bar{\mathcal{U}} = \begin{pmatrix} |U_x|^2 + |U_y|^2 & U_x\bar{U}_z - U_y\bar{U}_x \\ \bar{U}_x U_z - \bar{U}_y U_x & |U_z|^2 + |U_x|^2 \end{pmatrix}. \tag{2.162}$$

Since $\mathcal{U}\bar{\mathcal{U}}$ is a Hermitian matrix, we can use the $U(2)$ transformation, which has the form of a similarity transformtion on $\mathcal{U}\bar{\mathcal{U}}$ in Eq. (2.161), to diagonalise it. Then one can set the off-diagonal elements of $\mathcal{U}\bar{\mathcal{U}}$ to zero. When this is done, we obtain one condition for the action of the $U(2)$ transformation upon the quantities $\{U_x, U_y, U_z\}$, since the off-diagonal elements of $\mathcal{U}\bar{\mathcal{U}}$ are related by complex conjugation. This condition is given by:

$$U_x\bar{U}_z = U_y\bar{U}_x. \tag{2.163}$$

The only consistent solution to the condition Eq. (2.163) is $U_x = 0$. If $U_y = 0$ or $U_z = 0$, one necessarily generates additional conditions on U_y and U_z, which is inconsistent with their status as free parameters in the $U(N)$ two-instanton solution given by Eqs. (2.131,2.132,2.135,2.136). This procedure is not a standard one, we note. The ADHM construction does not fix the residual $U(k)$ symmetry by itself, and deducing which parameters are to be removed with these auxiliary transformations is a matter of choice. Hence our choice of $U(2)$ transformation for this instanton solution is not unique.

The $U(N)$ ADHM Three-Instanton Constraints

The $U(N)$ ADHM constraints for $k = 3$ present new difficulties not present in the $k = 2$ constraints. Attempts at straightforwardly generalizing the above method used to solve the $k = 2$ constraints to the $k = 3$ case fail because of the larger number of constraints and the appearance of many more bilinear terms. The coupled nature of the $k = 3$ constraints makes it extremely difficult to locate and exploit any linearity which may assist in solving the system of equations. As in the $k = 2$ case, there does not seem to be any underlying principle which could be applied to the constraints in order to solve them.

Here we state the $U(N)$ $k = 3$ constraints as derived from the ADHM construction. These were obtained by setting $k = 3$ in the ADHM constraints given in Eqs. (2.75,2.76). In this case we define the following quantities for convenience:

$$\bar{w}_1 w_2 \equiv \begin{pmatrix} U_1 & U_2 & U_3 \\ U_4 & U_5 & U_6 \\ U_7 & U_8 & U_9 \end{pmatrix}, \tag{2.164}$$

$$\bar{w}_1 w_1 - \bar{w}_2 w_2 \equiv \begin{pmatrix} U_x & U_{t1} & U_{t2} \\ \bar{U}_{t1} & U_y & U_{t3} \\ \bar{U}_{t2} & \bar{U}_{t3} & U_z \end{pmatrix}. \tag{2.165}$$

and write out the explicit form of the complex submatrices $(a'_2 + ia'_1)$ and $(a'_4 + ia'_3)$ as:

$$(a'_2 + ia'_1) = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \tag{2.166}$$

$$(a'_2 + ia'_1) = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}. \tag{2.167}$$

The instanton centre of mass co-ordinates for $k = 3$ are denoted X_i, $i = 1, 2, 3$.

The explicit $U(N)$ ADHM constraints for $k = 3$ then consist of the following fifteen conditions:

$$\begin{aligned}
U_1 + U_5 + U_9 &= 0, \\
U_x + U_y + U_z &= 0, \\
U_1 + \bar{a}_2\alpha_2 + \bar{a}_3\alpha_3 - \bar{b}_1\beta_1 - \bar{c}_1\gamma_1 &= 0, \\
U_2 + \beta_1\bar{X}_3 - \bar{a}_2X_1 + \bar{a}_3\beta_3 - \bar{c}_2\gamma_1 &= 0, \\
U_3 + \gamma_1(\bar{X}_3 + \bar{X}_4) - \bar{a}_3(X_1 + X_2) + \bar{a}_2\gamma_2 - \bar{b}_3\beta_1 &= 0, \\
U_4 + \bar{b}_1X_1 - \bar{c}_1(X_1 + X_2) + \bar{b}_3\alpha_3 - \bar{c}_1\gamma_2 &= 0, \\
U_5 + \bar{b}_1\beta_1 + \bar{b}_3\beta_3 - \bar{a}_2\alpha_2 - \bar{c}_2\gamma_2 &= 0, \\
U_6 + \gamma_2\bar{X}_4 - \bar{b}_3X_2 + \bar{a}_2\gamma_2 - \bar{b}_3\beta_1 &= 0, \qquad (2.168) \\
U_7 + \bar{c}_1(X_1 + X_2) - \alpha_3(\bar{X}_3 + \bar{X}_4) + \bar{c}_2\alpha_2 - \bar{b}_1\beta_3 &= 0, \\
U_8 + \bar{c}_2X_2 - \beta_3\bar{X}_4 + \bar{c}_1\beta_1 - \bar{a}_2\alpha_3 &= 0, \\
U_x + |a_2|^2 + |a_3|^2 - |b_1|^2 - |c_1|^2 + |\beta_1|^2 + |\gamma_1|^2 - |\alpha_2|^2 - |\alpha_3|^2 &= 0, \\
U_y + |b_1|^2 + |b_3|^2 - |a_2|^2 - |c_2|^2 + |\alpha_2|^2 + |\gamma_2|^2 - |\beta_1|^2 - |\beta_3|^2 &= 0, \\
U_{t1} + b_1\bar{X}_3 - \bar{a}_2X_3 + \bar{a}_3b_3 - \bar{c}_2c_1 + \bar{\alpha}_2X_1 - \beta_1X_1 + \bar{\gamma}_2\gamma_1 - \bar{\alpha}_3\beta_3 &= 0,
\end{aligned}$$

$$\begin{aligned} U_{t2} + c_1(\bar{X}_3 + \bar{X}_4) - \bar{a}_3(X_3 + X_4) + \bar{a}_2 c_2 - \bar{b}_3 b_1 & \\ + \bar{\alpha}_3(X_1 + X_2) - \gamma_1(\bar{X}_1 + \bar{X}_2) + \bar{\beta}_3\beta_1 - \bar{\alpha}_2\gamma_2 &= 0, \end{aligned}$$

$$U_{t3} + c_2\bar{X}_4 - \bar{b}_3 X_4 + \bar{c}_1 b_1 - \bar{a}_3 a_2 + \bar{\beta}_3 X_2 - \gamma_2\bar{X}_2 + \bar{\alpha}_3\alpha_2 - \bar{\beta}_1\gamma_1 = 0.$$

The $U(N)$ $k = 3$ ADHM constraints in Eqs. (2.168) are clearly of a much greater complexity than the $k = 2$ constraints. Any linearity present in these conditions is difficult to exploit due to the presence of the bilinear terms in all but the first two constraints given in Eqs. (2.168). Also, the greater number of constraints makes any manipulations and substitutions very lengthy and laborious to complete. These conditions are opaque and there does not appear to be any apparent method of solution which bears any resemblance to the method of solution for the $k = 2$ case. However, the constraints in Eqs. (2.168) do exhibit a pattern of terms similar to the $k = 2$ case.

The $U(N)$ ADHM Constraints for $k \geq 4$

We now describe an observation concerning the present scheme of identification of physical instanton parameters (instanton collective co-ordinates) and the $U(N)$ ADHM constraints for topological charge $k \geq 4$. The general $U(N)$ ADHM two-instanton solution was obtained above by adopting the identification of the instanton centre of mass co-ordinates with the diagonal elements of the submatrices $(a'_2 + ia'_1)$ and $(a'_4 + ia'_3)$ of the ADHM matrix a. The off-diagonal elements of the submatrices $(a'_2 + ia'_1)$ and $(a'_4 + ia'_3)$ were then eliminated using the ADHM constraints, resulting in a reasonable, but not unique, interpretation of the collective co-ordinates present in the matrix a which physically describe the instanton solution. The remaining collective co-ordinates of the instanton were formed from the elements of the other submatrices $w_{\dot{\alpha}}$ present in a. A number of elements of the submatrices $w_{\dot{\alpha}}$ are also eliminated by the ADHM constraints. Since some of the ADHM constraints are independent of the elements of the submatrices a'_n, and thus of the submatrices $(a'_2 + ia'_1)$ and $(a'_4 + ia'_3)$ within a'_n, these conditions could be used to eliminate variables in the submatrices $w_{\dot{\alpha}}$. This has consequences for the identification of instanton collective co-ordinates for topological charge $k \geq 4$.

We argue that the scheme for identifying the instanton collective co-ordinates above, which was applied to the $k = 1$ and $k = 2$ cases does not apply in the cases with $k \geq 4$. More precisely, there are an insufficient number of constraints to eliminate all of the off-diagonal elements of the submatrices a'_n for $k \geq 4$ provided that all of the gauge orientations are taken to be within the submatrices $w_{\dot{\alpha}}$. This leaves physical instanton parameters amongst the off-diagonal elements of a'_n, which does not occur in the $k = 2$ case. Hence a new identification of the physical parameters of the instanton contained within the ADHM matrix a is required. If such an identification is not possible, this raises questions about the physical validity of the ADHM construction, at least for the gauge group $U(N)$.

Our argument is as follows. The ADHM construction for a $U(N)$ k-instanton configuration requires the explicit and general solution of $3k^2$ real constraints. These $3k^2$ constraints place conditions on the total number of real parameters $4Nk + 4k^2$ present within the ADHM matrix a. There remains the $U(k)$ residual symmetry, which removes a further k^2 real parameters, leaving $4Nk$ real independent parameters for the description of the instanton. In the scheme above for the identification of the physical instanton collective co-ordinates, there are $4k(k-1)$ real parameters to be eliminated from the submatrices a'_n. These are the off-diagonal elements of a'_n, and once these are eliminated, the diagonal elements of a'_n can be interpreted as the centre of mass spacetime co-ordinates of the instanton. Of the $3k^2$ ADHM constraints, two constraints always remain independent of the elements of a'_n. These are given by the trace over the instanton number indices of each of the ADHM constraints in Eq. (2.75) and Eq. (2.76). Hence there are $3k^2 - 2$ ADHM constraints which can be used to eliminate the off-diagonal elements of a'_n or other elements in the submatrices $w_{\dot{\alpha}}$. However, if there are more ADHM constraints than there are off-diagonal elements within a'_n, then the aforementioned interpretation of the parameters contained in a cannot be made. This statement can be expressed quantitatively as the following inequality condition on the topological charge k of the instanton configuration:

$$3k^2 - 2 \ \geq 4k(k-1), \tag{2.170}$$

which implies that:

$$(4-k)k \ \geq \ 2. \tag{2.171}$$

The inequality in Eq. (2.171) implies that for $k \geq 4$ the conventional interpretation of the parameters in the ADHM matrix a cannot be made. The $U(k)$ residual symmetry cannot be used to rectify this inequality, since if the k^2 real parameters are not removed from the submatrices $w_{\dot{\alpha}}$, as is conventional (and used for the $k = 2$ case), then there must be k^2 parameters removed from the submatrices a'_n, or the removal of k^2 parameters must be split across both the submatrices $w_{\dot{\alpha}}$ and a'_n. Thus the global $U(k)$ symmetry transformations cannot be used to enable one to make the conventional interpretation the parameters of a.

We speculate that this statement has two possible implications. Firstly, $U(N)$ ADHM k-instanton configurations with $k \geq 4$ perhaps do require that some of the off-diagonal parameters of the submatrices a'_n are indeed to be identified as physical instanton parameters. However, in this case the completely clustered limit would leave more than $4k$ real parameters within the submatrices a'_n, making an identification of the centre of mass coordinates of the each of the individual constituent one-instanton solutions problematic. Secondly, there is the heretical possibility that the $U(N)$ ADHM construction breaks down for topological charge $k \geq 4$. Since no reasonable physical interpretation of the remaining free parameters can be made, the construction for $k \geq 4$ may not yield physical configurations. We have not yet investigated this apparent difficulty in identifying the physical parameters for the $Sp(N)$ ADHM construction when $k \geq 4$. In Subsection 2.3.3 we describe and consider the $Sp(N)$ ADHM construction, with emphasis upon the $k = 2$ and $k = 3$ cases.

2.3.3 $Sp(N)$ ADHM Multi-Instantons

The ADHM construction for instantons with symplectic gauge group $Sp(N)$ is useful as it gives a parameterization for instanton configurations with gauge group $SU(2)$ which is more economical and simple to use than the $SU(N)$ formalism. This is made possible by the existence of the isomorphism $Sp(1) \simeq SU(2)$. The gauge group $SU(2)$ is commonly chosen as it is the most simple choice of gauge group for Yang–Mills gauge theories.

Here we describe two approaches to the construction of $Sp(N)$ ADHM k-instantons. The first uses embedding the gauge group $Sp(N)$ into a larger $SU(2N)$ gauge group, and adapting the previous $SU(N)$ ADHM formalism to describe the $Sp(N)$ instantons. This is the approach taken in [224]. Other approaches to the construction of the $SU(N)$ and

$Sp(N)$ formalisms begin with larger $O(N)$ or $Sp(N)$ groups and embed the required gauge groups into these [22, 33].

The second approach uses a formalism which is intrinsically suited to the $Sp(N)$ gauge group. For this we follow the construction developed in [33]. In this approach one begins with a quaternionic formalism which requires no embedding into larger gauge groups. This construction requires a smaller number of variables and constraints than the modified $SU(N)$ formalism in the first approach. Hence, for simplicity and clarity, we adopt this second approach in our descriptions of the ADHM constraints for $Sp(N)$ instantons. The original $Sp(N)$ formalism is also the formalism employed in the majority of literature on $Sp(N)$ ADHM instantons [22, 23, 33].

In this subsection the known $Sp(N)$ three-instanton solutions are described [22, 23]. These solutions are special solutions of the $Sp(N)$ ADHM three-instanton constraints, and are the amongst the only known explicit instanton configurations of topological charge greater than $k = 2$ (see also [95], and in particular [96] for special (ADHM) instanton solutions of higher topological charge). We compare the three-instanton solutions given in [22, 23] and speculate on their relevance to the general solution of the $Sp(N)$ ADHM three-instanton constraints. In this subsection we also describe some conjectures regarding the form and properties of the general solution of the $Sp(N)$ ADHM three-instanton constraints.

The $Sp(N) \subset SU(2N)$ ADHM Construction

To derive the $Sp(N)$ ADHM construction from the $U(N)$ or $SU(N)$, the embedding $Sp(N) \subset SU(2N)$ can be used. (For the $SO(N)$ construction, one can use the embedding $SO(N) \subset SU(N)$.) By imposing appropriate reality conditions on the ADHM construction for $SU(2N)$ instantons of topological charge k, the gauge field v_m can be ensured to be valued in the $sp(N)$ subalgebra of the $su(2N)$ Lie algebra. To state the reality conditions we define the symplectic transpose t, which acts upon $Sp(N)$ group indices. For a column vector v, this transpose acts as $v^t = v^T J^T$, where T denotes the conventional matrix transpose. The matrix J is a $2N \times 2N$ sympletic matrix, given by:

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \qquad (2.172)$$

In the adjoint representation of $Sp(N)$, symplectic matrices are Hermitian matrices which are anti-symmetric under the transpose t, with dimension $(2N+1)N$. Hermitian matrices which are symmetric under t belong to the anti-symmetric representation of $Sp(N)$, and have dimension $(2N-1)N$.

The reality conditions on the $SU(2N)$ ADHM construction which give the $Sp(N)$ ADHM construction are then:

$$\bar{w}^{\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}}(w_{\dot{\beta}})^t, \qquad (2.173)$$

$$(a'_{\alpha\dot{\alpha}})^t = a'_{\alpha\dot{\alpha}}, \qquad (2.174)$$

where $\epsilon^{\dot{\alpha}\dot{\beta}}$ is the totally anti-symmetric rank two tensor defined in Appendix A, and J is the symplectic matrix defined in Eq. (2.172).

The reality conditions Eqs. (2.173,2.174) are invariant under the residual symmetry group attached to the $Sp(N)$ construction. For $Sp(N)$, this is the subgroup $O(k)$ of the $SU(2N)$ residual symmetry group $U(k)$. The Hermitian matrices a'_n are now also symmetric under the symplectic transpose t, and hence real and symmetric under the action of the $O(k)$ auxiliary group.

The $SU(N)$ ADHM constraints Eqs. (2.72,2.73) are adjoint-valued under $O(k)$. In conjunction with the reality conditions Eqs. (2.173,2.174), the $SU(2N)$ gauge field v_m generated by the ADHM construction will then be anti-symmetric under t. Since v_m is Hermitian, this gauge field also exists in the adjoint representation of $Sp(N)$. Hence the modified $SU(2N)$ ADHM construction can also be used to describe $Sp(N)$ k-instantons, and all of the previously given formulae and identities of the $U(N)$ ADHM construction hold for this instanton configuration. To be more precise, the gauge field v_m now takes values in the $sp(N)$ subalgebra of the $su(2N)$ algebra. This modification of the $SU(N)$ ADHM construction can also be applied to describe $Sp(N)$ fermion fields in the background of ADHM instantons.

The dimension of the modified $SU(2N)$ moduli space is equivalent to the dimension of the $Sp(N)$ moduli space. Including the action of the reality conditions Eqs. (2.173,2.174), one has $4Nk$ real parameters present in the submatrix w, and $2k(k+1)$ real parameters in the submatrices a'_n. The $SU(2N)$ ADHM constraints remove $3k^2$ real parameters, and the $O(k)$ residual symmetry group removes a further $\frac{1}{2}k(k+1)$ real parameters. This leaves a total of $4(N+1)k$ real parameters for the description of the $Sp(N)$ k-instanton. This is in agreement with the parameter counting for the $Sp(N)$ formalism below, which also gives the dimension of the $Sp(N)$ k-instanton moduli space as $4(N+1)k$, which is the same result the Atiyah–Singer index theorem gives upon applying it as in Subsection 2.2.2.

We now describe the $Sp(1)$ ADHM construction in this formalism, which is used in applications involving instantons in the theories with gauge group $SU(2)$ through the isomorphism $SU(2) \simeq Sp(1)$. The reality conditions Eqs. (2.173,2.174) for $Sp(1)$ are given explicitly by:

$$w^*_{ui\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}} J_{uv} w_{vi\dot{\beta}}, \tag{2.175}$$

$$(a'_n)_{ij} = (a'_n)_{ji}. \tag{2.176}$$

These conditions imply that a'_n are real symmetric $k \times k$ matrices, and that the submatrices w_i can be expressed as quaternions, $w_{i\alpha\dot{\alpha}} = w_{in}\sigma_{n\alpha\dot{\alpha}}$ and $w_i^{\dot{\alpha}\alpha} = w_{in}\bar{\sigma}_n^{\dot{\alpha}\alpha}$. Consequently, the ADHM matrix $a_{\dot{\alpha}}$ can be written as a quaternion:

$$a_{\dot{\alpha}} = \begin{pmatrix} w_{\alpha\dot{\alpha}} \\ a'_{\alpha\dot{\alpha}} \end{pmatrix}, \tag{2.177}$$

where the gauge group indices $u = 1, 2$ have now been written as Weyl indices $\alpha = 1, 2$. The submatrix $w_{\alpha\dot{\alpha}}$ is now a quaternion which satisfies the following relation in this specific $N = 2$ scheme:

$$w^*_{ui\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}}\epsilon^{\alpha\beta} w_{i\beta\dot{\beta}}. \tag{2.178}$$

The $Sp(1)$ and $SU(2)$ formalisms are isomorphic, but the formalism for the sympletic gauge groups has an advantage over the intrinsic $SU(2)$ formalism in that there are fewer variables and constraints in the $Sp(1)$ formalism than there are for that of $SU(2)$. When $N = 2$, the $SU(2)$ ADHM k-instanton construction has $4k(2+k)$ real parameters in $a_{\dot{\alpha}}$ on which $3k^2$ constraints and k^2 residual symmetries are imposed. For the $Sp(1)$ formalism, there are $2k(3+k)$ real parameters in $a_{\dot{\alpha}\alpha}$ on which $3k(k-1)/2$ constraints and

$k(k-1)/2$ residual symmetries are imposed. In both cases the number of real physical parameters is $8k$, as is required, but it is apparent that the $Sp(1)$ formalism presents a number of constraints almost a factor of two fewer than that for the $SU(2)$ formalism. For example, when $k = 1$, there are no constraints present in the $Sp(1)$ construction, whereas in the $SU(2)$ construction there are three. The $Sp(1)$ formalism also involves fixing fewer residual symmetries, which is a non-trivial task for $k > 1$, as can be seen from the $U(2)$ transformation for the $U(N)$ two-instanton solution above.

The $Sp(N)$ ADHM Construction

The explicit canonical form of the $Sp(N)$ and $U(N)$ ADHM constructions were first given in [33]. Here we follow the treatment of [33] for the $Sp(N)$ ADHM construction. We now denote the rank of the symplectic group by N, as there is now no embedding in a special unitary group. The symplectic group $Sp(N)$, in its compact and real (i.e. gauge group) form, can be described as the group of $N \times N$ matrices with quaternionic elements. Therefore, the ADHM construction for the symplectic groups may be given in terms of matrices of quaternions. A brief description of the properties of quaternions and the conventions we employ for quaternions are given in Appendix B. In particular, we note that quaternions are not commutative objects, and therefore the ordering of quaternionic variables in the following paragraphs is important.

The canonical form of the ADHM matrices a and b in the intrinsic $Sp(N)$ formalism can be written as:

$$a_{\lambda i} = \begin{pmatrix} w_{ui} \\ r_{ij} \end{pmatrix}, \quad b_{\lambda i} = \begin{pmatrix} 0_{[N+k]\times[N+k]} \\ -1_{[k]\times[k]} \end{pmatrix}, \tag{2.179}$$

where $r = r_{[k]\times[k]}$ and $w = w_{[N]\times[k]}$ are matrices of quaternions. The matrix indices in Eq. (2.179) have the same definitions and ranges as the indices in Eq. (2.46) for the $U(N)$ formalism. The $Sp(N)$ ADHM k-instanton constraints then assume the simple form:

$$r_{ij} = r_{ji}, \tag{2.180}$$

$$\bar{a}_{\lambda i} a_{\lambda j} = 0, \quad i \neq j. \tag{2.181}$$

The canonical form Eq. (2.179) are invariant under the transformation by the residual symmetry group $O(k)$, which acts in the same way upon a and b as in Eq. (2.133).

The $Sp(N)$ ADHM Two-Instanton

As there are no constraints for $k = 1$ in the $Sp(N)$ construction, the first ADHM constraints for symplectic gauge groups appear for $k = 2$. These constraints were first explicitly and generally solved in [22]. The ADHM matrix $a_{\dot{\alpha}\alpha}$ in this case has the form:

$$a_{\dot{\alpha}\alpha} = \begin{pmatrix} w_1 & w_2 \\ r_{11} & r_{12} \\ r_{12} & r_{22} \end{pmatrix}. \tag{2.182}$$

In terms of the quaternion variables contained in $a_{\dot{\alpha}\alpha}$, the $Sp(N)$ $k = 2$ ADHM constraints can be derived in terms of the canonical forms Eq. (2.179) from:

$$\bar{a}a = \begin{pmatrix} \bar{w}_1 w_1 + \bar{r}_{11}r_{11} + \bar{r}_{12}r_{12} & \bar{w}_1 w_2 + \bar{r}_{11}r_{12} + \bar{r}_{12}r_{22} \\ \bar{w}_2 w_1 + \bar{r}_{12}r_{11} + \bar{r}_{22}r_{12} & \bar{w}_2 w_2 + \bar{r}_{12}r_{12} + \bar{r}_{22}r_{22} \end{pmatrix}, \tag{2.183}$$

from which the $i \neq j$, or off-diagonal, elements give the two $k = 2$ ADHM constraints:

$$\bar{w}_1 w_2 + \bar{r}_{11}r_{12} + \bar{r}_{12}r_{11} = 0, \tag{2.184}$$

$$\bar{w}_2 w_1 + \bar{r}_{12}r_{11} + \bar{r}_{11}r_{12} = 0. \tag{2.185}$$

Note that the conditions in Eqs. (2.184,2.185) are the Hermitian conjugates of one another, so that there is actually only one constraint for the $Sp(N)$ ADHM two-instanton:

$$\bar{w}_1 w_2 + \bar{r}_{11}r_{12} + \bar{r}_{12}r_{11} = 0. \tag{2.186}$$

An approach to solving these constraints is to eliminate the off-diagonal elements, leaving a clear physical interpretation for the on-diagonal elements r_{11} and r_{22} as forming the instanton centre of mass co-ordinates. The constraint Eq. (2.186) is linear in the off-diagonal quaternion element r_{12}; however, it is also linear in its Hermitian conjugate $\bar{r}_{12}$. Therefore, to solve the constraint Eq. (2.186) we must also use its Hermitian conjugate Eq. (2.185). Subtracting Eq. (2.185) from Eq. (2.186) gives:

$$(\bar{w}_1 w_2 - \bar{w}_2 w_1) + \bar{X}_1 r_{12} - \bar{r}_{12} X_1 = 0, \tag{2.187}$$

where we have defined the quantity X_1:

$$X_1 \equiv r_{11} - r_{22}. \tag{2.188}$$

The constraint Eq. (2.187) has a unique general solution in which r_{12} is eliminated, but this solution cannot be formally derived by a series of algebraic operations. Instead, the general solution to Eq. (2.187), as given in [22], is essentially as an ansatz, which by observation solves Eq. (2.187) for the dependent variable r_{12}. The general explicit solution of the $Sp(N)$ two-instanton ADHM constraint in Eq. (2.186) is then:

$$r_{12} = \frac{1}{|X_1|^2} X_1 [U_1 + \Sigma], \tag{2.189}$$

where Σ is an arbitrary real number, and U_1 is the anti-Hermitian quaternion defined by:

$$U_1 \equiv \bar{w}_2 w_1 - \bar{w}_1 w_2. \tag{2.190}$$

The appearance of the term Σ in the solution Eq. (2.189) arises since the constraint Eq. (2.187) permits it. This is because an arbitrary real number may be added to the term U_1, which is a manifestly anti-Hermitian quaternion, in the solution for r_{12}, without affecting the constraint Eq. (2.187). At this point it becomes clear that the solution Eq. (2.189) has has not been derived formally, but rather has been derived by observation. However, this observation was completely sufficient to arrive at the correct exact solution of the constraints for this case.

An identification of the physical collective co-ordinates of the $Sp(N)$ two-instanton solution can be made as follows. The variable X_1 represents the relative separation of the two constituent one-instanton centre of mass positions, r_{11} and r_{22}. In the limit of large separation, which is the completely clustered limit, one has:

$$\lim_{|X_1| \to \infty} r_{12} \to 0, \tag{2.191}$$

so that the off-diagonal element r_{12} vanishes and the ADHM submatrix r becomes diagonal, and the matrix a for the $Sp(N)$ two-instanton can be explicitly written as the direct sum of two distinct $Sp(N)$ one-instanton configurations:

$$\lim_{|X_1| \to \infty} a \to \begin{pmatrix} w_1 & w_2 \\ r_{11} & 0 \\ 0 & r_{22} \end{pmatrix} = \begin{pmatrix} w_1 & 0 \\ r_{11} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & w_2 \\ 0 & 0 \\ 0 & r_{22} \end{pmatrix}. \tag{2.192}$$

The submatrix r is then proportional to the quaternionic spacetime co-ordinate multipled by b, which appears in the formula for $\Delta(x) = a + bx$. The scale sizes ρ_1, ρ_2 of the constituent $Sp(N)$ one-instantons will be related to the diagonal elements in the matrix $(\bar{a}a)$ given in Eq. (2.183). The gauge iso-orientations, which specify the orientation of the $Sp(1)$ solution in $Sp(N)$ group space will be contained in the anti-Hermitian quaternion U_1. We also note that the only clustering limit for the $k = 2$ solution above is the completely clustered limit in Eq. (2.192).

The explicit $Sp(N)$ two-instanton gauge field can be constructed using formulae similar to those for the $U(N)$ ADHM construction given in Subsection 2.3.1. However, we shall not give this here but note that for $Sp(1)$, any physical (gauge invariant) quantities derived using the $Sp(1)$ two-instanton gauge field shall be equivalent to those derived from the $SU(2)$ two-instanton gauge field which can be constructed from the $U(N)$ two-instanton solution given in Eqs. (2.127–2.132).

The $Sp(N)$ ADHM Three-Instanton

We now turn to the case of the $Sp(N)$ ADHM constraints with topological charge $k = 3$. As stated above, the $Sp(N)$ two-instanton constraints in Eqs. (2.184,2.185) are linear in the dependent variable. Although the method of solution is not a formal one, and indeed there is no discernible method for solution in that case other than an observation leading to an ansatz, the $k = 2$ constraints are readily solved. For the $k = 3$ constraints, formulating a suitable ansatz, from observation or otherwise, is greatly impeded by the form of the constraints, which are somewhat opaque. The $Sp(N)$ ADHM matrix a for $k = 3$ has the form:

$$a = \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & r_{12} & r_{13} \\ r_{12} & r_{22} & r_{23} \\ r_{13} & r_{23} & r_{33} \end{pmatrix} . \qquad (2.193)$$

As for the two-instanton case, the $Sp(N)$ ADHM three-instanton constraints can be derived from the matrix $\bar{a}a$, following Eqs. (2.180,2.181). The elements of $\bar{a}a$ which have

$i \neq j$, that is, the off-diagonal elements, give the $k = 3$ ADHM constraints as the following set of six conditions:

$$\begin{aligned}
\bar{w}_1 w_2 + \bar{r}_{11} r_{12} + \bar{r}_{12} r_{22} + \bar{r}_{13} r_{23} &= 0, && (2.194)\\
\bar{w}_1 w_3 + \bar{r}_{11} r_{13} + \bar{r}_{12} r_{23} + \bar{r}_{13} r_{33} &= 0, && (2.195)\\
\bar{w}_2 w_1 + \bar{r}_{12} r_{11} + \bar{r}_{22} r_{12} + \bar{r}_{23} r_{13} &= 0, && (2.196)\\
\bar{w}_2 w_3 + \bar{r}_{12} r_{13} + \bar{r}_{22} r_{23} + \bar{r}_{23} r_{33} &= 0, && (2.197)\\
\bar{w}_3 w_1 + \bar{r}_{13} r_{11} + \bar{r}_{23} r_{12} + \bar{r}_{33} r_{13} &= 0, && (2.198)\\
\bar{w}_3 w_2 + \bar{r}_{13} r_{12} + \bar{r}_{23} r_{22} + \bar{r}_{33} r_{23} &= 0. && (2.199)
\end{aligned}$$

Three of the six conditions in Eqs. (2.194–2.199) are related to one another by Hermitian conjugation, in a precisely similar way in which the $k = 2$ constraints Eqs. (2.184,2.185) are related to each other. Hence there are only three distinct $Sp(N)$ three-instanton constraints, which we choose to write as:

$$\begin{aligned}
\bar{w}_1 w_2 + \bar{r}_{11} r_{12} + \bar{r}_{12} r_{22} + \bar{r}_{13} r_{23} &= 0, && (2.200)\\
\bar{w}_1 w_3 + \bar{r}_{11} r_{13} + \bar{r}_{12} r_{23} + \bar{r}_{13} r_{33} &= 0, && (2.201)\\
\bar{w}_2 w_3 + \bar{r}_{12} r_{13} + \bar{r}_{22} r_{23} + \bar{r}_{23} r_{33} &= 0. && (2.202)
\end{aligned}$$

The $k = 3$ ADHM constraints in Eqs. (2.200–2.202) differ markedly from the $k = 2$ ADHM constraints in Eqs. (2.184,2.185). Firstly, there are three constraints for $k = 3$ and one for $k = 2$. The $k = 3$ constraints relate nine quaternions compared to the four in the $k = 2$ constraints. Each of the three $k = 3$ constraints contain a bilinear term involving the quaternions r_{ij}, $i \neq j$, which are absent for $k = 2$. Hence the $k = 3$ constraints are non-linear and consist of a set of coupled, simultaneous bilinear quaternionic conditions. In constrast, the $k = 2$ constraints are a set of simultaneous linear quaternionic conditions.

The three $k = 3$ constraints in Eqs. (2.200–2.202) have been solved in [22, 23] in special circumstances for $Sp(1)$. By fixing the $O(k) = O(3)$ symmetry and eliminating the complex part of one of the quaternion variables in a, the $k = 3$ $Sp(1)$ ADHM constraints are simplified and are amenable to solution. However, the number of collective co-ordinates which these solutions possess is then too few to constitute a general solution of the $k = 3$

$Sp(1)$ ADHM constraints. These solutions are therefore not described by the $Sp(N)$ extended instanton moduli space. We now state these solutions of the $Sp(1)$ ADHM three-instanton constraints in Eqs. (2.200–2.200).

The $k = 3$ solution given by Christ et. al [22] uses the residual $O(3)$ symmetry to set the complex part of the quaternion w_1 to zero. Thus the quaternion $w_1 = (w_{10}, w_{11}, w_{12}, w_{13})$ is set to the real parameter $w_1 = (w_{10}, 0, 0, 0) \in \mathbb{R}$. This removes three real parameters. Then the $Sp(1) \simeq SU(2)$ gauge symmetry of the solution is used to set the real parts of the three quaternions r_{ij}, $i \neq j$ to zero. This removes nine real parameters. After imposing the $k = 3$ constraints, which also remove nine real parameters, there remain twenty-one real parameters. The $k = 3$ $Sp(1)$ constraints are solved in [22] by eliminating the quaternions w_i as follows:

$$w_{10} = \frac{|\vec{W}_1 \times \vec{W}_3|}{\left[\vec{W}_1 \cdot (\vec{W}_2 \times \vec{W}_3)\right]^{1/2}} \in \mathbb{R}, \tag{2.203}$$

$$w_2 = w_{10} \frac{(\vec{W}_3 \times \vec{W}_2) \cdot (\vec{W}_3 \times \vec{W}_1)}{|\vec{W}_2 \times \vec{W}_3|^2} + i\vec{\tau} \cdot \frac{1}{w_{10}} \vec{W}_3, \tag{2.204}$$

$$w_2 = w_{10} \frac{(\vec{W}_3 \times \vec{W}_2) \cdot (\vec{W}_2 \times \vec{W}_1)}{|\vec{W}_2 \times \vec{W}_3|^2} - i\vec{\tau} \cdot \frac{1}{w_{10}} \vec{W}_2, \tag{2.205}$$

where the complex vectors W_i are defined by:

$$\vec{W}_i = \frac{i}{4} \epsilon_{ijk} \mathrm{tr}_2 \left[\vec{\tau}(\bar{r}_{ii} - \bar{r}_{jj}) r_{ij} + \sum_{l=1}^{3} \bar{r}_{li} r_{lj} \right], \tag{2.206}$$

where $\vec{\tau}$ are the standard Pauli matrices given in Appendix A.

The $k = 3$ solution in Eqs. (2.203–2.205) is elaborate and contains a square root of a complex quantity in the denominator of the real parameter w_{10}. This solution is claimed to be generalized to the gauge group $Sp(N)$ by using the global $Sp(N)$ gauge transformations and $O(k)$ residual symmetry to set the complex parts of the Nk quaternions w_{iu} to zero. This solution is expressed in complicated terms and to our knowledge has not been used to explicitly derive the $SU(2)$ three-instanton gauge field. Furthermore, a physical interpretation of the solution, and indentification of the three-instanton collective co-ordinates, is not forthcoming and may not be possible.

An alternative solution to the $Sp(N)$ ADHM three-instanton constraints was proposed by Korepin and Shatashvili [23]. These authors use the residual $O(3)$ symmetry to set the complex part of the quaternion parameter r_{13} to zero. Then r_{13} is set to the real number $r_{13} = R_{13} \in \mathbb{R}$. The solution to the $k = 3$ $Sp(N)$ ADHM constraints is obtained by exploiting the form of the constraints in a non-trivial manner. This ensures that the solution is given in terms of rational functions of a subset of parameters within the matrix a. The $k = 3$ $Sp(N)$ solution in [23] is given by:

$$w_1 = -[\bar{r}_{11}r_{12} + \bar{r}_{12}r_{22} + \bar{r}_{13}r_{23}]\frac{1}{r_{11}}, \tag{2.207}$$

$$\bar{r}_{23} = \left[\frac{1}{r_{13}}(\bar{r}_{22}r_{11} + \bar{r}_{12}\bar{w}_1)r_{23} - \bar{r}_{22}r_{12} - \bar{r}_{12}r_{11}\right]\frac{1}{\left(r_{22} - \dfrac{r_{12}r_{23}}{\bar{r}_{13}}\right)}, \tag{2.208}$$

$$\bar{r}_{33} = -\frac{1}{r_{13}}[\bar{r}_{22}r_{11} + \bar{r}_{23}r_{12} + \bar{r}_{12}\bar{w}_1]. \tag{2.209}$$

The solution in Eqs. (2.207–2.209) of the $k = 3$ $Sp(N)$ ADHM constraints has twenty-one instanton collective co-ordinates, a feature which it has in common with the solution previously described in Eqs. (2.203–2.205). Since the two solutions given in Eqs. (2.203–2.205) [22] and Eqs. (2.207–2.209) [23] have the same number of instanton parameters, they should only differ from one another by no more than the permitted local $O(3)$ gauge transformations. However, they actually differ by more than this transformation. This is because the $k = 3$ solution in Eqs. (2.203–2.205) [22] is given in terms of irrational functions of the elements of a, whereas the $k = 3$ solution determined in Eqs. (2.207–2.209) [23] has a rational parameterization. A rational parameterization for the solution of the $Sp(N)$ $k = 3$ constraints precludes roots of quaternionic quantities being taken. If roots of quaternions, which can be realized as complex matrices, are taken, then this introduces the possibility that there are complex phases attached to the results of the roots. Since the instanton gauge field v_m is Hermitian (but anti-Hermitian in our conventions), one would not expect v_m to depend on variables which have complex phases, as this implies that the gauge field strength v_{mn} can be complex. This is unphysical and contradictory, as v_{mn} is defined to be Hermitian and thus real, as it is observable. Therefore, if the $k = 3$ constraints possess an irrational solution, any roots in the parameterization must have real arguments.

In [23], the $Sp(N)$ three-instanton instanton gauge field is constructed, albeit in an indirect way, and the result is a rational function of the independent parameters in the solution Eqs. (2.207–2.209). Hence the $Sp(N)$ three-instanton gauge field so constructed cannot be complex, and, as is physically required, will be a real, Hermitian $N \times N$ matrix. For the purposes of instanton calculus, and progress in solving the ADHM constraints, one requires the general solution of the $Sp(N)$ three-instanton constraints in Eqs. (2.200–2.202). The general $Sp(N)$ three-instanton solution shall have twenty-four independent parameters. A possible approach to the construction of the general solution is to 'un-fix' the $O(3)$ residual symmetry transformations which were used to obtain special solutions [22, 23] of the $Sp(N)$ three-instanton constraints. Seeking a rational parameterization, we attempted this procedure for the solution given in Eqs. (2.207–2.209) [23]. However, difficulty was encountered in this attempt. Specifically, the twenty-one parameter solution given in Eqs. (2.207–2.209) appears to have eliminated variables which have a preferred physical interpretation. This observation also applies to the alternative irrational solution in Eqs. (2.203–2.205). The elimination of parameters usually identified as physical is possible because the $O(3)$ residual symmetry used in the solution method has already removed parameters usually taken to be physical, namely the complex part of the quaternion w_1. The set of quaternions $\{w_{ui}\}$ will contain some parameters which are gauge group orientations included in the physical parameter count for the extended $Sp(N)$ k-instanton moduli space. Therefore, implementing the residual symmetry transformation before variables are eliminated through the constraints implies that some of the gauge group orientations may have already been removed. For this reason, there exists no clear physical interpretation of either of the three-instanton solutions given in [22, 23]. This is at least the case in terms of the conventional identifications made for unitary group ADHM multi-instantons, already described in Subsection 2.3.2 . The proposed twenty-one parameter three-instanton solutions can be considered $SU(2)$ ADHM three-instanton solutions when $N = 1$. Therefore the identification of physical parameters made for unitary multi-instantons should apply to these solutions in that case. It is not clear that this can be done for the above three-instanton solutions.

Alternatively, one may begin with the $k = 3$ $Sp(N)$ ADHM constraints and determine the general solution. We discuss this possibility in the next paragraph.

Properties of $Sp(N)$ ADHM Three-Instantons

The $k = 3$ $Sp(N)$ ADHM constraints in Eqs. (2.200–2.202) present the problem of a set of coupled non-linear quaternion equations which are bilinear in the dependent variables to be eliminated. In this case, for a straightforward physical interpretation similar to that of the $k = 2$ case, the dependent variables are the off-diagonal quaternion elements r_{12}, r_{13} and r_{23} of the matrix a. It may be possible to exploit similarities or an analogy between the $k = 2$ and $k = 3$ ADHM constraints in seeking a $k = 3$ solution.

The $k = 2$ constraints are contained in the $k = 3$ constraints in the following sense. If the quaternions r_{13} and r_{23} are set to zero, then the $k = 3$ $Sp(N)$ ADHM constraints in Eqs. (2.200–2.202) reduce to:

$$\bar{w}_1 w_2 + \bar{r}_{11} r_{12} + \bar{r}_{12} r_{22} = 0, \tag{2.210}$$

$$\bar{w}_1 w_3 = 0, \tag{2.211}$$

$$\bar{w}_2 w_3 = 0. \tag{2.212}$$

If one takes w_1 and w_2 to be non-zero, then the solution to Eqs. (2.211,2.212) is $w_3 = 0$. The matrix a of Eq. (2.179) then assumes the form:

$$a = \begin{pmatrix} w_1 & w_2 & 0 \\ r_{11} & r_{12} & 0 \\ r_{12} & r_{22} & 0 \\ 0 & 0 & r_{33} \end{pmatrix}, \tag{2.213}$$

which is the $k = 2$ $Sp(N)$ ADHM matrix a directly summed with a 3×3 matrix with a single entry r_{33} in the lowest right hand column component. This property does not appear to be useful in determining the general $k = 3$ $Sp(N)$ solution, however.

The $k = 3$ matrix a also has a number of clustering limits. In the first clustering limit, which is a partially clustered limit, one expects that $Sp(N)$ three-instanton will decompose into an $Sp(N)$ one-instanton and an $Sp(N)$ two-instanton. In seeking a general solution to the $Sp(N)$ three-instanton constraints which has a valid physical interpretation, we anticipate that the off-diagonal elements of the submatrix r should be eliminated. This then leaves the on-diagonal elements of r to be identified as the instanton centre of mass co-ordinates, and their differences as the relative instanton separations. In this approach, for the partially clustered limit in which $k = 3$ decomposes to a direct sum of single $k = 1$ and $k = 2$ instantons, there are three possible ways in which this can occur. These limits are obtained by simultaneously taking a pair of the off-diagonal elements of r to be zero:

$$a \rightarrow \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & 0 & r_{13} \\ 0 & r_{22} & 0 \\ r_{13} & 0 & r_{33} \end{pmatrix}, \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & r_{12} & 0 \\ r_{12} & r_{22} & 0 \\ 0 & 0 & r_{33} \end{pmatrix}, \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & 0 & 0 \\ 0 & r_{22} & r_{23} \\ 0 & r_{23} & r_{33} \end{pmatrix}. \tag{2.214}$$

A further complication of the partially clustered limit for the $k = 3$ instanton is the existence of overlapping configurations in the partially clustered limit, where two $Sp(N)$ two-instantons may arise which are coupled in such a way as to produce an $Sp(N)$ one-instanton and an $Sp(N)$ two-instanton. There are three of these overlapping partially clustered limits possible for the $Sp(N)$ three-instanton. These are given by:

$$a \rightarrow \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & r_{12} & 0 \\ r_{12} & r_{22} & r_{23} \\ 0 & r_{23} & r_{33} \end{pmatrix}, \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & 0 & r_{13} \\ 0 & r_{22} & r_{23} \\ r_{13} & r_{23} & r_{33} \end{pmatrix}, \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & r_{12} & r_{13} \\ r_{12} & r_{22} & 0 \\ r_{13} & 0 & r_{33} \end{pmatrix}. \tag{2.215}$$

In the limits given in Eq. (2.215), the $Sp(N)$ three-instanton constraints in Eqs. (2.200–2.202) assume the form of two $Sp(N)$ $k = 2$ constraints, as in Eq. (2.186), for the two non-zero off-diagonal quaternions remaining in a, plus a constraint which couples the same two off-diagonal quaternions. The second clustering limit is the completely clustered limit, in which the $Sp(N)$ three-instanton will decompose into three $Sp(N)$ one-instantons. Again, we consider this limit in the scheme stated above, in which the instanton collective coordinates within the submatrix r are given by its diagonal elements. In this limit, which is obtained by simultaneously sending all of the off-diagonal elements of r to zero, the ADHM three-instanton matrix a then becomes:

$$a \rightarrow \begin{pmatrix} w_1 & w_2 & w_3 \\ r_{11} & 0 & 0 \\ 0 & r_{22} & 0 \\ 0 & 0 & r_{33} \end{pmatrix}. \tag{2.216}$$

These limits might be used as physical, rather than mathematical, constraints, on the form of the general solution to the $k = 3$ $Sp(N)$ ADHM constraints. However, it is difficult to see how they could be used to further constrain the existing constraints.

Using the same physical interpretation of the elements of the $Sp(N)$ ADHM matrix a, one can count the number of possible clustering limits which an $Sp(N)$ ADHM k-instanton configuration can possess. For example, there shall always be one and only one completely clustered limit, which for $k = 3$ is given by Eq. (2.216). Furthermore, there will be $\frac{1}{2}k(k-1)$ partially clustered limits; for $k = 3$ these are given by the three limits in Eq. (2.214). The number of overlapping clustering limits, which for $k = 3$ are the three limits in Eq. (2.215), varies according to k. The total number of clustering limits for each $Sp(N)$ ADHM k-instanton configuration (within this scheme of physical interpretation of the instanton parameters) can be calculated combinatorically. Since the ADHM quaternionic submatrix r_{ij} is symmetric, there are a set of $n = \frac{1}{2}k(k-1)$ objects which can become zero. Of these, the number $r = 1, 2, \ldots, k$ can be simultaneously zero. The clustering limits can be classified according to this number. For $r = 1$, the limit is that of partial clustering; for $2 \leq r \leq (k-1)$, the limit is either that of partial clustering or overlapping clustering, depending on k, and for $r = k$, one has the unique completely clustered limit. Using the standard formula for the number of combinations ${}^nC_r$ of r objects selected from a set of n (without replacement), the total number of clustering limits is then given by the sum:

$$\mathfrak{C} = \begin{pmatrix} n \\ 1 \end{pmatrix} + \begin{pmatrix} n \\ 2 \end{pmatrix} + \cdots + \begin{pmatrix} n \\ k \end{pmatrix} = \sum_{r=1}^{n} \frac{n!}{(n-r)!r!}. \tag{2.217}$$

where $n = \frac{1}{2}k(k-1)$ and we have used the standard definition of ${}^nC_r$:

$${}^nC_r \equiv \begin{pmatrix} n \\ r \end{pmatrix} = \frac{n!}{(n-r)!r!}. \tag{2.218}$$

The formula for the total number of possible clustering limits of the $Sp(N)$ ADHM k-instanton configuration given by $\mathfrak{C}$ in Eq. (2.217) yields $\mathfrak{C} = 1$ for $k = 2$ and $\mathfrak{C} = 7$ for $k = 3$. The total number of clustering limits for these values of k agree with the possible $k = 2$ and $k = 3$ clustering limits given above respectively in Eq. (2.192) and Eqs. (2.214,2.215,2.216).

We now present a list of some properties which we conjecture the general $Sp(N)$ three-instanton solution should possess. In making these conjectures, we assume that the physical identification of collective co-ordinates within the ADHM matrix a follow the interpretation used for the general $Sp(N)$ two-instanton solution. We conjecture that:

1. As the solution is expected to have a rational parameterization, we can write the off-diagonal elements of the submatrix r as an irreducible fraction. That is, we conjecture that the general solution has the form:

$$r_{ij} = r_{ji} = f_i(X) = \frac{p_i(X)}{q_i(X)}, \quad i \neq j, \tag{2.219}$$

where p_i and q_i are polynomials in $X = \{X_i\}$ with no common factors, and where the variables $\{X_i\}$ are the relative instanton separations given by the differences between the on-diagonal elements of r:

$$X_1 \equiv r_{11} - r_{22}, \tag{2.220}$$

$$X_2 \equiv r_{11} - r_{33}, \tag{2.221}$$

$$X_3 \equiv r_{22} - r_{33}. \tag{2.222}$$

It follows that the ADHM matrix $U(X)$ in the $Sp(N)$ construction is a rational function of X and the spacetime co-ordinate x, since only the combination $(x - X)$ can appear in the solution. We also conjecture that a rational parameterization of the ADHM matrix $\Delta(X)$ must exist for each value of k in the ADHM construction if the general k-instanton gauge field v_m is to be real and Hermitian. This conjecture is supported by the explicit form of the general $Sp(N)$ two-instanton and the general $U(N)$ two-instanton solutions given in this chapter.

2. Given the form of the dependent variables r_{ij} in conjecture (1.), using the physical interpretation of the elements of r implies that the one must have:

$$\mathrm{Deg}\,(q_i) > \mathrm{Deg}\,(p_i), \tag{2.223}$$

where $\mathrm{Deg}(h)$ is the degree of the polynomial quaternion $h(X)$ in X. We conjecture this property in order that the correct clustering limits are satisfied in the limit $X \to \infty$, and that the cluster decompositions Eqs. (2.214,2.216) exist. Heuristically, we note that the inequality Eq. (2.223) assumes the form of an equality for both the $Sp(N)$ and $U(N)$ two-instanton solutions, namely:

$$\mathrm{Deg}\,(q_i) = \mathrm{Deg}\,(p_i) + 1. \tag{2.224}$$

Whether Eq. (2.224) holds for the $Sp(N)$ or $U(N)$ three-instantons or other multi-instanton configurations remains unknown. This conjecture is in analogy with the work of Donaldson on the monopole moduli space [287], in which it was shown that the monopole moduli space is diffeomorphic to the space of rational functions which vanish at infinity. The requirement that instanton configurations possess rational parameterizations and obey the dilute instanton gas limit will constrain the instanton moduli space in a similar way, although general exact solutions of the ADHM constraints are expected to automatically satisfy these requirements. Our conjectures in Eqs. (2.223,2.224) regarding the degrees of the polynomials p_i and q_i in Eq. (2.219) are in accord with the dilute instanton gas limit, and ensure that the off-diagonal elements of the submatrix r_{ij} vanish at infinity.

3. The functions $f_i(X)$ given in Eq. (2.219) are globally odd functions of the relative instanton separations $X = \{X_i\}$. This conjecture is based upon an heuristic observation concerning the $Sp(N)$ ADHM constraints for $k \geq 2$. If one reflects all of the relative instanton separations $\{X_i\}$ appearing in the $Sp(N)$ ADHM constraints about the origin, then the ADHM constraints must not generate new constraints which further constrain the existing implicitly defined instanton configuration. If such a transformation does generate new conditions, then the ADHM constraints will become trivial and readily solved, leading to a trivial instanton solution. Therefore, the reflection $\{X_i\} \to \{-X_i\}$ should not produce new constraints for the instanton configuration. Given the form of the $Sp(N)$ ADHM constraints, the only way to avoid generating new constraints is to assume that the quaternions r_{ij} are also reflected under the transformation $\{X_i\} \to \{-X_i\}$. This is because the quaternion elements $w_{\dot{\alpha}iu}$ are independent of the relative instanton separations $\{X_i\}$. That is, when $\{X_i\} \to \{-X_i\}$, then one must simultaneously have $r_{ij}(X) \to -r_{ij}(X)$. If this holds, then the ADHM constraints are invariant under the transformation $\{X_i\} \to \{-X_i\}$, and no new constraints are generated by acting upon the ADHM constraints with this reflection. It follows that $f_i(X)$ must be a globally odd function in $\{X_i\}$ since we have defined $r_{ij} = r_{ji} = f_i(X)$ in Eq. (2.219). Explicitly, the functions $f_i(X)$ then have the property of being odd functions of the relative

instanton separations $\{X_i\}$:

$$f_i(\{-X_i\}) + f_i(\{X_i\}) = 0. \qquad (2.225)$$

This property is exhibited in the explicit form of the general $Sp(N)$ two-instanton and the general $U(N)$ two-instanton solutions given in this chapter.

4. Conjecture (3.) in conjunction with conjecture (1.) then implies the following three corollaries. Firstly, if the polynomial $p_i(X)$ is a globally even function of $\{X_i\}$, then $q_i(X)$ is a globally odd function of $\{X_i\}$. Conversely, if $p_i(X)$ is a globally odd function of $\{X_i\}$, then $q_i(X)$ is a globally even function of $\{X_i\}$. The explicit form of the general $Sp(N)$ and $U(N)$ two-instanton solutions include the polynomial $q_i(X)$ as an even function of $\{X_i\}$. Secondly, since the quaternions $\{U_i\}$ (formed from the quaternions $w_{\dot{\alpha}iu}$) appearing in the $Sp(N)$ $k=2$ and $k=3$ constraints are independent of the parameters $\{X_i\}$, then they are globally even functions of $\{X_i\}$. Therefore, additive terms of the form:

$$\sum_i \left(X_i \pm \tfrac{1}{2}U_i\right), \quad \sum_{i,j,k} \left(X_i X_j X_k \pm \tfrac{1}{2}U_i\right), \ \ldots$$

are forbidden to appear in the polynomials $p_i(X)$ and $q_i(X)$, since such terms are neither odd nor even functions of $\{X_i\}$. The factor of 1/2 multiplying the $\{U_i\}$ terms is a normalization factor, and the dots indicate terms of higher degree in $\{X_i\}$. However, these terms are permitted in $p_i(X)$ and $q_i(X)$ when multiplied by an odd factor of $\{X_i\}$, such as:

$$\sum_i X_i \left(\tfrac{1}{2}U_i \pm \alpha_i\right), \quad \sum_{i,j,k} X_i X_j X_k \left(\tfrac{1}{2}U_i \pm \alpha_i\right),$$

where $\{\alpha_i\}$ are a set of real constants independent of both $\{U_i\}$ and $\{X_i\}$. These permitted terms are manifestly odd functions in $\{X_i\}$, and may therefore appear in the polynomials $p_i(X)$, if the $p_i(X)$ are taken to be odd functions of $\{X_i\}$. Although these terms are permitted, they are neither obligatory nor necessary, and may not appear in, for example, the general $Sp(N)$ three-instanton solution. Note, though, that terms of the permitted form do appear in the general $Sp(N)$ two-instanton solution. Thirdly, we remark that conjecture (2.) and (3.) may be equivalent to one another, or imply one another.

5. Lastly, we conjecture that the $Sp(N)$ ADHM constraints can only be solved at the level of quaternions. Unlike the $U(N)$ ADHM construction, for which the method of solution for the $k = 2$ case involved splitting a complex constraint into real and imaginary parts, a similar procedure was not found to be useful for the $Sp(N)$ three-instanton constraints. When the constraints in Eqs. (2.200–2.202) are written out in their component forms, any method of solution for individual components of quaternions becomes unviable. Associated with this there is also the difficulty in reconstructing quaternions from their separate components, as detailed in Appendix B. Therefore, we conjecture that the $Sp(N)$ ADHM constraints only admit solutions in terms of quaternions, and not their components. This is the case for the $Sp(N)$ two-instanton constraints, and given the same scheme of identification of physical parameters, we conjecture that this shall also be the case for the $Sp(N)$ three-instanton constraints. This difficulty in using the $Sp(N)$ ADHM constraints in their component form may also be related to the lack of a formal method of solution for the $Sp(N)$ two-instanton constraints given in Eq. (2.186). There may not exist any formal method of solution for the $Sp(N)$ ADHM constraints with any topological charge k, leaving one to determine an ansatz for the $Sp(N)$ three-instanton solution which will solve the constraints in Eqs. (2.200–2.202).

We also note that some work towards $Sp(N)$ four-instanton solutions has been performed. These include an attempt to construct an $Sp(N)$ four-instanton following the method used to obtain the special $Sp(N)$ three instanton [23] reproduced above in Eqs. (2.207–2.209). Unfortunately, Korepin and Shatashvili [23] found that the method used for the three-instanton case did not generalize to the four-instanton case, which appears to be a generic feature of the ADHM constraints. That is, methods of solution for the ADHM constraints of a given topological charge k do not extend to those with topological charge greater than or equal to $k + 1$. This indicates that there are no underlying principles which might be used to solve the k-instanton ADHM constraints through some iterative or algorithmic procedure. This perspective on the ADHM constraints is reinforced by the work of Inozemtsev [24], in which an attempt to determine a special $Sp(N)$ four-instanton solution using the special $Sp(N)$ three-instanton solution of Christ et. al [22], reproduced in Eqs. (2.203–2.205) [22] above, is made. This work is intriguing, but is not successful, which again implies that the ADHM constraints do not admit an iterative procedure for explicitly determining multi-instantons of higher topological charge from those of lower topological charge.

It is hoped that the ADHM construction of instantons can be of future use in the study of non-perturbative quantum field theory, perhaps in a modified form. Otherwise, due to its inherent difficulties, the ADHM construction has returned diminishing amounts of information regarding instantons over time. Since its inception, the ADHM construction has proven itself the most simple and effective method for the explicit construction of instanton configurations, both special and general. Other methods which can be used to solve the self-dual Yang–Mills field equations, which include [86, 85, 84, 88, 87], (we note Refs. [86, 88] actually pre-date the ADHM construction), have not been as productive as the ADHM construction. We note that following the success of the ADHM construction for self-dual and anti-self-dual gauge fields, studies of non-self-dual gauge fields were initiated [89, 100, 101, 102, 103], including attempts at treating these gauge fields in terms of algebraic geometry [100], as in the ADHM construction. The most simple example of a non-self-dual gauge field is the combination of a self-dual and an anti-self-dual gauge field [91]. The resulting instanton configuration has zero topological charge but remains a non-trivial gauge field configuration.

The ADHM construction has also been adapted to describe all monopole gauge field solutions of classical Yang–Mills gauge theory. Monopole configurations can be described by an infinite dimensional version of the ADHM construction, proposed by Nahm [104], and consequently known as the Nahm construction. The Nahm construction [104] has been used to generate general monopole solutions with arbitrary classical gauge group, which have subsequently been used in semi-classical calculations (for example, in [144]), and also investigated for generalization to the non-self-dual case [106]. Monopoles also possess topological charge, and the most general explicit solution of the Nahm monopole constraints is known only for the two-monopole, analogous to the ADHM instanton case [105]. For a review of Yang–Mills monopoles (and other classical Yang–Mills field configurations), see for example [75]. For a review of current research on monopoles in similar directions, one may consult [107].

At present the ADHM construction has been used most successfully in instanton calculations in quantum supersymmetric gauge theories. (The ADHM construction has a formalism which is also manifestly supersymmetry invariant [108].) Due to the properties of supersymmetric gauge theories, instanton effects can be calculated exactly, and interest in the ADHM construction has been renewed. In Chapter 3 below we describe global supersymmetry and supersymmetric gauge theories.

Chapter 3

Supersymmetric Gauge Theories

3.1 Introduction

Symmetry is an important guiding principle in theoretical physics. In particle physics, symmetries assume a particularly prominent rôle: gauge symmetry, for instance, is at the foundation of particle physics. Supersymmetry is a special symmetry which connects fermions and bosons via the unification of internal and external symmetries. It is at time of writing a purely theoretical construction, with no complete experimental evidence for its existence. Despite this, supersymmetry provides a fascinating theoretical laboratory for quantum field theory. Through supersymmetry, many simplifications regarding quantum field theory can be made, which for instance permits one to calculate quantities exactly in some supersymmetric field theories. It is noteworthy that supersymmetry is the only known extension of present quantum field theory consistent with the known properties of quantum fields [121]. This fact immediately suggests that any physics not included within the Standard Model of particle physics is likely to be supersymmetric. This includes gravitation, and attempts at formulating field theories which unify the known fundamental forces often include supersymmetry.

Supersymmetry is present either locally or globally. The case of local supersymmetry, which gives rise to supersymmetric gravity or supergravity theories, will not be considered here. Globally supersymmetric classical field theories will be our focus in this chapter, and all supersymmetry described in this thesis shall be global supersymmetry.

A classical supersymmetric field theory is essentially a classical field theory with a particular matter content. Symmetries not existing in the matter content of ordinary classical field theories arise due to the supersymmetry in these theories, and these can be exploited.

In Section 3.2 we outline the basic formalism of supersymmetry using superfields. We follow this with the construction of specific classical supersymmetric gauge theories and their properties. In Section 3.3, we described the simplest classical supersymmetric gauge theory, which possesses $\mathcal{N} = 1$ supersymmetry; exact results in these theories, and a form of duality connecting $\mathcal{N} = 1$ supersymmetric gauge theories with different gauge groups, known as Seiberg duality, will be described in Chapter 4. In Section 3.4, $\mathcal{N} = 2$ classical supersymmetric gauge theory is reviewed. Exact results in classical and quantum field theories with $\mathcal{N} = 2$ will be described in Chapter 5. In Section 3.5 the simple case of $\mathcal{N} = 3$ classical supersymmetric gauge theory is briefly described. We conclude this chapter with a brief review of $\mathcal{N} = 4$ classical supersymmetric gauge theory in Section 3.6; these theories shall be described further in Chapter 4.

3.2 Supersymmetry

In this section a brief overview of supersymmetric field theory is given. We make use is made of the extensive reviews [120, 121], the standard texts [122, 126, 127] and the reviews [190, 191]. Other useful reviews, particularly for supersymmetry phenomenology, include [123, 124, 125]. We employ the notation and conventions of [122, 190, 191] and also use [123].

Supersymmetry originated for many reasons [121]. The primary reason was the search for extending known particle physics and the unification of the known fundamental forces. The S-matrix in quantum field theory describes particle interactions and is used in calculating the outcome of particle collisions and scattering events. A symmetry of the S-matrix corresponds to a symmetry transformation of the theory in which (asymptotic) single and multi-particle states are interchanged. The known symmetries of the four-dimensional S-matrix have been classified and are Poincaré invariance, internal global symmetries and discrete symmetries. Poincaré invariance combines rotational invariance and Lorentz invariance. Internal global symmetries are symmetries such as conservation of quantum numbers (e.g. charges, isospin, baryon number), and the generators form a Lie algebra. Discrete symmetries are C, P, T, denoting charge conjugation, parity reversal and time reversal symmerty, respectively. The Coleman-Mandula theorem [109] states that, under certain assumptions, these three sets of symmetries are the only possible symmetries of the S-matrix. The Coleman-Mandula theorem places restrictions on the possible form of new models of particle physics, ruling out, for instance, the $SU(6)$ grand unified theory [109]. One of the assumptions made in formulating the Coleman-Mandula theorem is that the symmetry algebra of the S-matrix involves commutators only. If this assumption is weakened to include anti-commutators also, then the Coleman-Mandula theorem is circumvented, and other symmetries of the S-matrix, corresponding to new physics, are possible.

Supersymmetry utilizes both commuting and anti-commuting symmetry generators, and so the Coleman-Mandula theorem does not apply. The anti-commuting symmetry generators exist in spinor representations of the Lorentz group. This makes supersymmetry a symmetry in which internal symmetries (possessing scalar generators) and external symmetries (Poincaré spacetime symmetries) are mixed together via the anti-commuting spinor generators. In this sense, supersymmetry unifies internal and external symmetries. Furthermore, supersymmetry can be viewed as extending the Poincaré spacetime symmetry group by the inclusion of spinor generators, and can exist in spacetime dimensions greater than four. The Coleman-Mandula theorem, however, is restricted to four dimensional spacetime.

A powerful result pertaining to supersymmetry was obtained by Haag, Lopuszański and Sohnius [116]. They showed that supersymmetry is in general the only additional symmetry of the S-matrix permitted by weakening the assumptions of the Coleman-Mandula theorem described above. (However, we note that conformally invariant theories will also possess conformal symmetry of the S-matrix.) This implies that supersymmetry is the only possible extension of the known spacetime symmetries consistent with existing particle physics. If the assumptions of the Coleman-Mandula theorem are weakened further, other physically relevant symmetries could be possible, but none have been found at present [123]. These facts motivate supersymmetry as the most likely extension of known particle physics, from a theoretical perspective.

In a supersymmetric theory, there exist the same number of bosonic and fermionic degrees of freedom. The bosonic and fermionic states will have the same mass and the same external quantum numbers. The fermionic states are paired up with an equal number of new bosonic states, and the original bosonic states of the theory are paired up with an equal number of fermionic states. The newly introduced particle states which are paired with the original field content of the theory are known as superpartner particles.

We note the original papers on supersymmetry in [110, 111, 112], and also the original papers on supersymmetric fields and gauge theories in [113, 114, 115, 117, 118, 119]. The mathematical work in [128] is also noted. General gauge theory references include those in [278, 279, 280, 281, 282]. The local version of supersymmetry, known as supergravity, is reviewed in [129]. Other related works on supersymmetry and its phenomenological implications (such as dynamical supersymmetry breaking) which we note can be found in [130, 131, 132].

In this chapter we work in four-dimensional Minkowski spacetime with metric $\eta_{ab} = \mathrm{diag}\,(-1,1,1,1)$, unless stated otherwise, as detailed in Appendix A. Our conventions for supersymmetry follow those given in [122].

3.2.1 Global Supersymmetry

A supersymmetric theory is one in which the component fields of the theory form a representation of the supersymmetry algebra. The most general supersymmetry algebra is that which includes central charges, one statement of which is:

$$\{Q^A_\alpha, \bar{Q}_{\dot\beta B}\} = 2\sigma^m_{\alpha\dot\beta} P_m \delta^A_B, \tag{3.1}$$

$$\{Q^A_\alpha, Q^B_\beta\} = 2\sqrt{2}\epsilon_{\alpha\beta} Z^{AB}, \quad \{\bar{Q}_{\dot\alpha A}, \bar{Q}_{\dot\beta B}\} = 2\sqrt{2}\epsilon_{\dot\alpha\dot\beta} Z^*_{AB}, \tag{3.2}$$

$$\left[P_m, Q^A_\alpha\right] = \left[P_m, \bar{Q}_{\dot\alpha A}\right] = 0, \tag{3.3}$$

$$[P_m, P_n] = 0. \tag{3.4}$$

The Greek indices α,β,$\dot{\alpha}$,$\dot{\beta}$ specify two-component Weyl spinors and assume the values 1,2; the lower case Latin indices m,n are indices for the components of Lorentz four-vectors and run from 0 to 3; lastly, the upper case Latin indices A and B are the supersymmetry indices: these specify elements in the internal space and run from 1 to some integer $\mathcal{N} \geq 1$. Here Q_α^A and $\bar{Q}_{\dot{\beta}B}$ are anti-commuting operators which act on a Hilbert space and P_m is the energy-momentum operator. The structure of the supersymmetry algebra in Eqs. (3.1–3.4) is that of a $\mathbb{Z}_2$ graded Lie algebra. The most general four-dimensional supersymmetry algebra will, in addition to the above algebra, include commutators relating the supersymmetry generators to the Lorentz symmetry generator M_{mn}.

Theories with $\mathcal{N} = 1$ supersymmetry are often referred to as having simple, or non-extended, supersymmetry. Theories which have more than one supersymmetry, $\mathcal{N} > 1$, are described as possessing extended supersymmetry.

In order to treat the fermionic and bosonic states in a unified manner, the supersymmetry algebra must be represented without mass-shell conditions. This can be done by defining multiplets of component fields which transform under supersymmetric operations. The component fields belonging to a given multiplet will each have the same mass.

The supersymmetry algebra can be written in terms of commutators only by using the constant Grassmann-valued anti-commuting parameters, or constant spinors, θ^α,$\bar{\theta}_{\dot{\alpha}}$, which obey:

$$\{\theta^\alpha, \theta^\beta\} = \{\bar{\theta}_{\dot{\alpha}}, \bar{\theta}_{\dot{\beta}}\} = \{\theta^\alpha, \bar{\theta}_{\dot{\beta}}\} \tag{3.5}$$

$$\{\theta^\alpha, Q_\beta\} = [P_m, \theta^\alpha] = 0. \tag{3.6}$$

The supersymmetry algebra then becomes:

$$\left[\theta Q, \bar{\theta}\bar{Q}\right] = 2\theta\sigma^m\bar{\theta}P_m, \tag{3.7}$$

$$[\theta Q, \theta Q] = \left[\bar{\theta}\bar{Q}, \bar{\theta}\bar{Q}\right] = 0, \tag{3.8}$$

$$[P^m, \theta Q] = \left[P^m, \bar{\theta}\bar{Q}\right] = 0, \tag{3.9}$$

in which the spinor summation conventions as given in [122] have been used. If the Grassmann-valued parameters θ^α,$\bar{\theta}_{\dot{\alpha}}$ were not constant and depended on spacetime, then the supersymmetry would be local supersymmetry, which we do not consider in this thesis.

3.2.2 Supersymmetry Constraints

In this subsection we describe a different formalism for the expression of supersymmetric field theories. In doing so, we retrace some of the original steps taken in formulating the Lagrangians for supersymmetric gauge theories [118, 119, 121], which will be described in Section 3.3. In Subsection 3.2.3, the Lie group elements of the $\mathcal{N}=1$ supersymmetry algebra in Eq. (3.19) are not gauge invariant. This motivates the introduction of a fermionic counterpart for the non-Abelian vector potential (or Yang–Mills potential) v_m, which is effectively the superpartner of v_m, and is known as the Yang–Mills spinor potential or superconnection, which we denote by $\mathcal{A}^i_\alpha$, where $i,j=1,\ldots,\mathcal{N}$. The supersymmetry covariant derivatives D_α and $\bar{D}_{\dot{\alpha}}$ defined in Eq. (3.19) can be made gauge invariant by the addition of the Yang–Mills spinor potential. In the general case of $\mathcal{N}$-extended supersymmetry, this invokes the following gauge invariant supersymmetry covariant derivatives [119]:

$$\mathcal{D}^i_\alpha = D^i_\alpha + \mathcal{A}^i_\alpha = \frac{\partial}{\partial\theta^\alpha_i} + i\sigma^m_{\alpha\dot{\alpha}}\partial_m\bar{\theta}^{\dot{\alpha}} + \mathcal{A}^i_\alpha, \tag{3.10}$$

$$\bar{\mathcal{D}}_{\dot{\alpha}} = \bar{D}_{\dot{\alpha}} + \bar{\mathcal{A}}_{\dot{\alpha}i} = -\frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}i}} - i\theta^\alpha_i\sigma^m_{\alpha\dot{\alpha}}\partial_m + \bar{\mathcal{A}}_{\dot{\alpha}i}. \tag{3.11}$$

To simplify notation, we now adopt the following shorthand notation $D^\alpha_i \equiv D_A$ and $\bar{D}_{\dot{\beta}i} \equiv \bar{D}_{\dot{B}}$. The commutator or anti-commutator of any two covariant derivatives, as defined in Eq. (3.10,3.11), yields an $\mathcal{N}$-supersymmetric field strength or curvature, referred to as the Yang–Mills field strengths or supercurvatures [119]:

$$\{\mathcal{D}_A, \mathcal{D}_B\} = v_{AB}, \quad \{\bar{\mathcal{D}}_{\dot{A}}, \bar{\mathcal{D}}_{\dot{B}}\} = \bar{v}_{\dot{A}\dot{B}}, \quad \{\mathcal{D}_A, \bar{\mathcal{D}}_{\dot{B}}\} = v_{A\dot{B}} - 2i\sigma^m_{A\dot{B}}\mathcal{D}_m, \tag{3.12}$$

$$[\mathcal{D}_m, \mathcal{D}_A] = v_{mA}, \quad [\mathcal{D}_m, \bar{\mathcal{D}}_{\dot{A}}] = \bar{v}_{m\dot{A}}, \quad [\mathcal{D}_m, \mathcal{D}_n] = v_{mn}, \tag{3.13}$$

where v_{mn} is the standard (non-supersymmetric) Yang–Mills gauge field strength, and the other field strengths are Lie-algebra valued supersymmetric fields. Henceforth, we refer to 'supersymmetric fields', meaning (classical) fields invariant under some number of supersymmetries, as 'superfields', as is conventional. This anticipates the superfield formalism which we described in Subsection 3.2.3. The six Yang–Mills superfield strengths in Eqs. (3.12,3.13) obey Bianchi identities [119], are gauge covariant and reside in the adjoint representation of the gauge group. Specific supersymmetric gauge theories can be expressed as a set of supersymmetry constraints which take the form of conditions on the Yang–Mills superfield strengths. The supersymmetry constraints should not give a flat theory, in which $v_{mn} = 0$, and should not reduce to equations of motion, otherwise the theory shall be trivial. These requirements on the constraints can be ensured by use of the Bianchi identities for the Yang–Mills superfield strengths.

The supersymmetry constraints for $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory, in which $v_{AB} = v_{\alpha\beta}$ are:

$$v_{\alpha\beta} = \bar{v}_{\dot{\alpha}\dot{\beta}} = v_{\alpha\dot{\beta}} = 0. \tag{3.14}$$

The $\mathcal{N} = 1$ supersymmetry constraints do not produce a flat theory because the superfield strength $v_{m\alpha}$ may contain an unconstrained spin $\frac{1}{2}$ field, which the constraints neither preclude nor determine [119]. The supersymmetry constraints in Eq. (3.14) lead to the the conditions Eq. (3.53,3.54) on $\mathcal{N} = 1$ vector superfields in Subsection 3.2.3, which constrain the form which the Lagrangian for $\mathcal{N} = 1$ supersymmetric Yang–Mills theory can assume. The supersymmetry constraints for $\mathcal{N} = 2$ Yang–Mills theory are the weaker conditions [118]:

$$\begin{aligned} v^{ij}_{\alpha\beta} + v^{ji}_{\beta\alpha} &= 0, \\ \bar{v}_{\dot{\alpha}i\dot{\beta}j} + \bar{v}_{\dot{\beta}i\dot{\alpha}j} &= 0, \\ v^{i}_{\alpha\dot{\beta}j} &= 0. \end{aligned} \tag{3.15}$$

Use of the Bianchi identities for the $\mathcal{N} = 1$ Yang–Mills superfields yield conditions on the $\mathcal{N} = 2$ supersymmetric Yang–Mill field strengths. These conditions indicate the suitability of the vector superfield $\mathcal{W}$ as the superfield with which to construct the $\mathcal{N} = 2$ supersymmetric Yang–Mills Lagrangian. The Lagrangian given in Eq. (3.81) of Section 3.4 emerges as a suitable Lagrangian for $\mathcal{N} = 2$ supersymmetric Yang–Mills theory.

Supersymmetry constraints provide a concise formulation of supersymmetric field theories and were originally used in the derivation of some supersymmetric gauge theories. The supersymmetry constraints can be solved in the cases of $\mathcal{N} = 1$ and $\mathcal{N} = 2$, and the solutions give the superfields of the theory in terms of component fields. Methods of solution for the constraints of $\mathcal{N} = 3$ supersymmetry have also been studied.

3.2.3 $\mathcal{N} = 1$ Supersymmetry and Superfields

The most simple supersymmetry is $\mathcal{N} = 1$ supersymmetry, when the supersymmetry algebra Eqs. (3.9) has only one supersymmetry generator ($A = B = 1$ in the supersymmetry algebra). Furthermore, the central charges Z_{AB} in the supersymmetry algebra Eqs. (3.1–3.4) vanish by antisymmetry.

To facilitate the construction of an $\mathcal{N} = 1$ supersymmetric field theory, the fields which appear in the theory are constrained to exist in sets, or multiplets [112]. The fields within multiplets are the component fields from which manifestly supersymmetric field theories can be constructed.

The multiplet of component fields is a set of fields $\{A, \psi, \ldots\}$ which transform under the infinitesimal transformation δ_θ as follows:

$$\delta_\theta A = (\xi Q + \bar{\theta}\bar{Q})A, \tag{3.16}$$

$$\delta_\theta \psi = (\xi Q + \bar{\theta}\bar{Q})\psi, \tag{3.17}$$

$$\ldots$$

These supersymmetry transformations map tensor fields into spinor fields and vice versa. Via these transformations, one can define the multiplets with specific properties. By constructing the appropriate multiplet of component fields and using the constraint of supersymmetry invariance, supersymmetric field theories can be formulated [113, 114].

Every representation of the supersymmetry algebra must contain an equal number of bosonic and fermionic states. Representations of the supersymmetry algebra can be explicitly constructed via the isomorphism between the algebra of the supersymmetry generators Q and algebras of fermionic and bosonic creation and annihilation operators, which act on a Clifford (algebra) vacuum. In this way all of the possible particle states of a representation can be determined. However, in this approach on-mass-shell and off-mass-shell cases must be treated separately: when $P^2 = -M^2$ or $P^2 = 0$, where M is the mass of the supersymmetry multiplet, one must define a separate set of operators to construct the particle states (these cases correspond, respectively, to the fermionic and bosonic operator cases).

An alternative and more intuitive method for constructing supersymmetric Lagrangians is the superfield formalism [115], which we shall adopt henceforth. Superfields are supersymmetric fields which are functions of the component fields, and the component fields contained within the superfield can be recovered by expanding the superfield in the anti-commuting parameters θ and $\bar{\theta}$. Superfields reside in superspace, the supersymmetric generalization of four dimensional spacetime.

The simplified supersymmetry algebra in Eq. (3.9) can be interpreted as an ordinary Lie algebra with anti-commuting elements. A Lie group can be associated with this algebra, and differential operators which generate translation in the parameters of group elements, and so relate group elements, are found to be:

$$Q_\alpha = \frac{\partial}{\partial\theta^\alpha} - i\sigma^m_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\partial_m, \quad \bar{Q}^{\dot{\alpha}} = \frac{\partial}{\partial\bar{\theta}_{\dot{\alpha}}} - i\theta^\alpha\sigma^m_{\alpha\dot{\beta}}\eta^{\dot{\beta}\dot{\alpha}}\partial_m, \tag{3.18}$$

$$D_\alpha = \frac{\partial}{\partial\theta^\alpha} + i\sigma^m_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\partial_m, \quad \bar{D}^{\dot{\alpha}} = -\frac{\partial}{\partial\bar{\theta}_{\dot{\alpha}}} - i\theta^\alpha\sigma^m_{\alpha\dot{\alpha}}\partial_m. \tag{3.19}$$

These linear differential operators satisfy the following anti-commutation relations:

$$\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\} = 2i\sigma^m_{\alpha\dot{\alpha}}\partial_m, \tag{3.20}$$

$$\{Q_\alpha, Q_\beta\} = \{\bar{Q}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = 0, \tag{3.21}$$

$$\{D_\alpha, \bar{D}_{\dot{\alpha}}\} = -2i\sigma^m_{\alpha\dot{\alpha}}\partial_m, \tag{3.22}$$

$$\{D_\alpha, D_\beta\} = \{\bar{D}_{\dot{\alpha}}, \bar{D}_{\dot{\beta}}\} = 0, \tag{3.23}$$

$$\{D_\alpha, Q_\beta\} = \{D_\alpha, \bar{Q}_{\dot{\beta}}\} = \{\bar{D}_{\dot{\alpha}}, Q_\beta\} = \{\bar{D}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = 0. \tag{3.24}$$

The group space spanned by the parameters $\{x, \theta, \bar{\theta}\}$ is referred to as $\mathcal{N} = 1$ superspace, the most simple supersymmetric generalization of four dimensional spacetime. The Grassmann-valued parameters $\{\theta, \bar{\theta}\}$ exist at each point in spacetime and effectively make the spacetime anticommuting. In general, $\mathcal{N}$-supersymmetric superspace shall have $4\mathcal{N}$ supersymmetry co-ordinates in addition to the four contained in x^m. Thus the $\mathcal{N} = 1$ superspace is eight dimensional. Furthermore, the supersymmetry generators $\bar{D}_{\dot{\alpha}}$ and D_α satisfy the supersymmetry algebra Eqs. (3.1–3.4) of Subsection 3.2.1 with the sign change $P_m \to -P_m$. In this operator representation of the supersymmetry generators, the momentum tensor P_m can also be written as a linear differential operator, $P_m = i\partial_m$, familiar from quantum mechanics.

The $\mathcal{N} = 1$ supersymmetry algebra possesses an internal Abelian symmetry algebra which gives rise to an additional internal global $U(1)$ symmetry. This symmetry is known as R-symmetry, and the generator R which generates this symmetry acts as follows on the $\mathcal{N} = 1$ supersymmetry generators:

$$[Q_\alpha, R] \;=\; Q_\alpha, \tag{3.25}$$

$$[\bar{Q}_{\dot{\alpha}}, R] \;=\; -\bar{Q}_{\dot{\alpha}}. \tag{3.26}$$

The $\mathcal{N} = 1$ supersymmetry generators $\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\}$ are said to have R-charge +1 and -1, respectively. The R-symmetry acts on the superfields of the $\mathcal{N} = 1$ supersymmetric gauge theory, and will resurface in Section 3.3.

The most general superfield on $\mathcal{N} = 1$ superspace, denoted $F(x, \theta, \bar{\theta})$, is given by a power series expansion in the Grassmann co-ordinates $\{\theta, \bar{\theta}\}$, whose coefficients are the component fields previously introduced. Following this, $F(x, \theta, \bar{\theta})$ has the form:

$$\begin{aligned} F(x,\theta,\bar{\theta}) \; &= f(x) + \theta\phi(x) + \bar{\theta}\bar{\chi}(x) \\ &+ \theta^2 m(x) + \bar{\theta}^2 n(x) + \theta\sigma^m\bar{\theta}v_m(x) \\ &+ \theta^2\bar{\theta}\bar{\sigma}(x) + \bar{\theta}^2\theta\phi(x) + \theta^2\bar{\theta}^2 d(x), \end{aligned} \tag{3.27}$$

in which higher order terms vanish due to the properties of θ and $\bar{\theta}$. Our conventions for products of supersymmetric co-ordinates are $\theta^\alpha\theta_\alpha = \theta^2$ and $\theta\sigma^m\theta = \theta^\alpha\sigma^m_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}$, following [122]. The field content of the superfield $F(x, \theta, \bar{\theta})$ consists of Weyl spinor, scalar and vector fields, a mixture of fermions and bosons. These fields form a representation of the supersymmetry algebra Eq. (3.1–3.4) of Subsection 3.2.1.

A general superfield $F(x, \theta, \bar{\theta})$ is defined to transform under a general supersymmetry transformation as:

$$\delta_\xi F(x,\theta,\bar{\theta}) \;\equiv\; (\xi Q + \bar{\xi}\bar{Q})F(x,\theta,\bar{\theta}), \tag{3.28}$$

where ξ_α and $\bar{\xi}_{\dot{\alpha}}$ are arbitrary Grassmann-valued parameters. The supersymmetry transformation laws for each component field appearing in $F(x, \theta, \bar{\theta})$ can be deduced by comparison of powers of θ,$\bar{\theta}$. Linear combinations of superfields are again superfields due to these transformation properties, and products of superfields are also superfields, since

Q and $\bar{Q}$ are linear differential operators. Superfields provide a linear representation of the supersymmetry algebra, and specific constraints on the general superfield $F(x, \theta, \bar{\theta})$ will yield specific supersymmetry representations. Given a multiplet of component fields, a superfield can always be constructed. A superfield may be constructed from a multiplet of component fields by acting on each of the component fields with the operator $\exp(\theta Q + \bar{\theta}\bar{Q})$. For example, given a field A, of general type, the corresponding superfield is given by:

$$F(x, \theta, \bar{\theta}) = e^{(\theta Q + \bar{\theta}\bar{Q})} A = A + \delta_\theta A + \cdots \tag{3.29}$$

Supersymmetric generalizations of any kind of field may be constructed via this method.

Chiral Superfields

The chiral superfield Φ is given by the most general solution to the constraint:

$$\bar{D}_{\dot{\alpha}} \Phi = 0, \tag{3.30}$$

which has the form:

$$\Phi = A(y) + \sqrt{2}\theta\psi(y) + \theta^2 F(y), \tag{3.31}$$

where a new variable $y^m = x^m + i\theta\sigma^m\bar{\theta}$ has been used. The chiral superfield Φ contains all of the component fields of the scalar multiplet: namely, the spinor A, the scalar ψ and the additional field F, and it can be seen that the superfield formalism is much more concise than the formalism which uses component fields alone. The field F within Φ is also known as the 'F-component' of Φ and transforms by a total derivative under $\mathcal{N} = 1$ supersymmetry transformations. Furthermore, the component field F is described as an 'auxiliary field' on account of it being determined (on-shell) algebraically by the remaining physical field content of Φ in the Lagrangian for the scalar multiplet. In general, auxiliary fields can be eliminated from a Lagrangian by using the Euler-Lagrange equations (equations of motion) of the theory. Thus auxiliary fields are described as being 'non-dynamical fields'; however, strictly they are fields which possess dynamics independent of the other fields in the Lagrangian. In this sense, with respect to the other (non-auxiliary) fields in the Lagrangian, they are not dynamical fields.

Expanding y in Eq. (3.31) yields:

$$\Phi = A(x)+i\theta\sigma^m\bar{\theta}\partial_m A(x)-\frac{1}{4}\theta^2\bar{\theta}^2\Box A(x)+\sqrt{2}\theta\phi(x)-\frac{i}{\sqrt{2}}\theta^2\partial_m\phi(x)\sigma^m\bar{\theta}+\theta^2F(x). \quad (3.32)$$

Any function of the variables $(y,\theta,\bar{\theta})$ is a chiral superfield, due to the following necessary and sufficient condition:

$$\bar{D}_{\dot{\alpha}}y^m = 0, \quad \bar{D}_{\dot{\alpha}}\theta^\beta = 0. \quad (3.33)$$

The operators D_α and $\bar{D}_{\dot{\alpha}}$ can be realized in terms of the variables $(y,\theta,\bar{\theta})$ as:

$$D_\alpha = \frac{\partial}{\partial\theta^\alpha}+2i\sigma^m_{\alpha\dot{\alpha}}\frac{\partial}{\partial y^m}, \quad (3.34)$$

$$\bar{D}_{\dot{\alpha}} = -\frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}}. \quad (3.35)$$

In a similar way, the anti-chiral superfield $\Phi^\dagger$ is the most general solution to the constraint:

$$D_\alpha\Phi^\dagger = 0, \quad (3.36)$$

and has the form:

$$\Phi^\dagger = A^\dagger(y^\dagger)+\sqrt{2}\bar{\theta}\bar{\psi}(y^\dagger)+\bar{\theta}^2F^\dagger(y^\dagger), \quad (3.37)$$

where $y^{m\dagger} = x^m - i\theta\sigma^m\bar{\theta}$ is the Hermitian conjugate of y^m; from this fact, the expansion of $\Phi^\dagger$ in terms of x and $\bar{\theta}$ will be the conjugate of Eq. (3.32). We note that the supersymmetry transformation law for Φ will give the individual transformation laws for each of the component fields $\{A,\psi,F\}$ of the scalar multiplet. As previously, any function of the variables $(y^\dagger,\bar{\theta})$ will also be an anti-chiral superfield. The general properties of superfields hold for Φ and $\Phi^\dagger$, with the exception that the product $\Phi\Phi^\dagger$ is not a scalar superfield. An arbitrary function of chiral superfields is also a chiral superfield:

$$\begin{aligned}\mathcal{W}(\Phi_i) &= \mathcal{W}(A_i+\sqrt{2}\theta\psi_i+\theta^2F_i)\\ &= \mathcal{W}(A_i)+\sqrt{2}\frac{\partial\mathcal{W}}{\partial A_i}\theta\psi_i+\theta^2\left(\frac{\partial\mathcal{W}}{\partial A_i}F_i-\frac{1}{2}\frac{\partial^2\mathcal{W}}{\partial A_i\partial A_j}\psi_i\psi_j\right). \quad (3.38)\end{aligned}$$

Such functions $\mathcal{W}$ are known as superpotentials and are holomorphic. The chiral superfields Φ and $\Phi^\dagger$ can be expressed in terms of the initial variables $(x,\theta,\bar{\theta})$ as follows:

$$\Phi = A(x)+i\theta\sigma^m\bar{\theta}\partial_m A-\frac{1}{4}\theta^2\bar{\theta}^2\Box A+\sqrt{2}\theta\psi(x)-\frac{i}{\sqrt{2}}\theta^2\partial_m\psi\sigma^m\bar{\theta}+\theta^2F(x), \quad (3.39)$$

$$\Phi^\dagger = A^\dagger(x)-i\theta\sigma^m\bar{\theta}\partial_m A^\dagger-\frac{1}{4}\theta^2\bar{\theta}^2\Box A^\dagger+\sqrt{2}\bar{\theta}\bar{\psi}(x)-\frac{i}{\sqrt{2}}\bar{\theta}^2\sigma^m\partial_m\bar{\psi}+\bar{\theta}^2F^\dagger(x)(3.40)$$

Chiral superfields obey the following transformation laws under the action of a local Abelian group:

$$\Phi_k \rightarrow \Phi'_k = e^{-it_k\lambda(x)}\Phi, \quad \bar{D}_{\dot{\alpha}}\lambda(x) = 0, \tag{3.41}$$

$$\Phi^\dagger_k \rightarrow \Phi'^\dagger_k = e^{it_k\bar{\lambda}(x)}\bar{\Phi}, \quad D_\alpha\bar{\lambda}(x) = 0, \tag{3.42}$$

in which t_k are the Abelian charges associated with each chiral superfield Φ_k.

The non-Abelian generalization of this transformation law for non-Abelian chiral superfields in the adjoint representation is:

$$\Phi \rightarrow \Phi' = e^{2ig\Lambda}\Phi e^{-2ig\Lambda}, \quad \Phi^\dagger \rightarrow \Phi^{\dagger\prime} = e^{2ig\Lambda^\dagger}\Phi^\dagger e^{-2ig\Lambda^\dagger}, \tag{3.43}$$

where g is the gauge coupling and Λ is the matrix defined by:

$$\Lambda_{ij} = T^a_{ij}\Lambda_a, \tag{3.44}$$

where the matrices T^a are the generators of the gauge group in the representation in which Φ resides:

$$\Phi = \Phi^a T^a. \tag{3.45}$$

When Φ is in a given representation of the gauge group, one can also specify $\mathrm{Tr}\, T^a T^b = k\delta^{ab}$ for $k > 0$ and the algebra $[T^a, T^b] = if^{abc}T^c$, where f^{abc} are the structure constants of the gauge group.

Vector Superfields

Vector superfields V are defined as superfields which obey the reality constraint:

$$V = V^\dagger. \tag{3.46}$$

This condition is solved in general by the following expression for V in terms of component fields:

$$\begin{aligned} V(x,\theta,\bar{\theta}) &= C(x) + i\theta\chi(x) - i\bar{\theta}\bar{\chi}(x) + \frac{i}{2}\theta^2[M(x) + iN(x)] - \frac{i}{2}\bar{\theta}^2[M(x) - iN(x)] \\ &\quad - \theta\sigma^m\bar{\theta}v_m(x) + i\theta^2\bar{\theta}\left[\bar{\lambda}(x) + \frac{i}{2}\bar{\sigma}^m\partial_m\chi(x)\right] \\ &\quad - i\bar{\theta}^2\theta\left[\lambda(x) + \frac{i}{2}\sigma^m\partial_m\bar{\chi}(x)\right] + \frac{1}{2}\theta^2\bar{\theta}^2\left[D(x) + \frac{1}{2}\Box C(x)\right]. \end{aligned} \tag{3.47}$$

In this specific component field expansion for $V(x, \theta, \bar{\theta})$, the fields $\{C, D, M, N, v_m\}$ must be Hermitian so that V qualifies as a vector superfield according to the constraint Eq. (3.46). The vector field v_m is the usual vector field of field theory, but the vector field v_m is not a gauge field until it has been gauged. From the form of the Hermitian combination of chiral and anti-chiral superfields $\Phi + \Phi^\dagger$, one can define supersymmetric gauge (or 'supergauge') transformations, the set of which forms an Abelian group:

$$V \to V + \Phi + \Phi^\dagger. \tag{3.48}$$

Under this Abelian gauge transformation, the component fields $\lambda(x)$ and $D(x)$ are gauge invariant, and the gauge can be fixed so that $C = \chi = M = N = 0$. This specific gauge is the Wess–Zumino or WZ gauge [113]. This choice of gauge breaks supersymmetry and is thus 'non-supersymmetric'. The Abelian gauge symmetry of v_m remains unfixed. We adopt the Wess–Zumino gauge hereon in this thesis.

In the Wess–Zumino gauge, the vector superfield V becomes:

$$V = -\theta\sigma^m\bar{\theta}v_m(x) + i\theta^2\bar{\theta}\bar{\lambda}(x) - i\bar{\theta}^2\theta\lambda(x) + \frac{1}{2}\theta^2\bar{\theta}^2 D(x). \tag{3.49}$$

It is notable that this form of V has the following properties:

$$V^2 = -\frac{1}{2}\theta^2\bar{\theta}^2 v_m v^m, \tag{3.50}$$

$$V^3 = 0. \tag{3.51}$$

The vector multiplet comprises a massless vector field v_m, the spinor field λ_α and the additional auxiliary field D. The component field D appearing in V, also known as the 'D-component' of the vector superfield V transforms by a total derivative under $\mathcal{N} = 1$ supersymmetry transformations.

Since the square of V is quadratic in the gauge field v_m, V may be considered as the supersymmetric generalization of the Yang–Mills gauge field strength (or vector potential) v_m. Continuing with this analogy leads to the definition of the supersymmetric field strength for the vector superfield V. The fields λ_α and $\bar{\lambda}_{\dot{\alpha}}$ possess the lowest dimension of the component fields in V, and they are also gauge invariant. The superfields W_α and $\bar{W}_{\dot{\alpha}}$ also possess λ_α and $\bar{\lambda}_{\dot{\alpha}}$ as their lowest dimensional component fields, where:

$$W_\alpha = -\frac{1}{4}\bar{D}\bar{D}D_\alpha V, \quad \bar{W}_{\dot{\alpha}} = -\frac{1}{4}DD\bar{D}_{\dot{\alpha}}V. \tag{3.52}$$

The superfields W and $\bar{W}$ are defined to be the Abelian field strengths for the vector superfield V. The superfields W and $\bar{W}$ are both chiral superfields:

$$\bar{D}_{\dot{\beta}} W_\alpha = 0, \quad D_\beta \bar{W}_{\dot{\alpha}} = 0, \tag{3.53}$$

and are also gauge invariant on account of the constraints $\bar{D}\Phi = D\Phi^\dagger = 0$. They also obey the constraint:

$$D^\alpha W_\alpha = \bar{D}_{\dot{\alpha}} \bar{W}^{\dot{\alpha}}. \tag{3.54}$$

As W_α is gauge invariant, it can be expanded in component fields in the Wess–Zumino gauge, and reads as:

$$W_\alpha = -i\lambda_\alpha(y) + \theta_\alpha D - \frac{i}{2}(\sigma^m \bar{\sigma}^n \theta)_\alpha v_{mn} + \theta^2(\sigma^m \partial_m \bar{\lambda})_\alpha, \tag{3.55}$$

where $v_{mn} = \partial_m v_n - \partial_n v_m$ is the Abelian gauge fields strength tensor. That the superfield W can be considered as the supersymmetric field strength of the vector superfield V is apparent from the $\theta\theta$ component of the product $W^\alpha W_\alpha$:

$$W^\alpha W_\alpha|_{\theta\theta} = -\frac{1}{4}\bar{D}\bar{D}W^\alpha D_\alpha V|_{\theta\theta} = -2i\lambda\sigma^m\partial_m\bar{\lambda} - \frac{1}{2}v^{mn}v_{mn} + D^2 + \frac{1}{2}iv^{mn*}v_{mn}, \tag{3.56}$$

in which ${}^*v_{mn}$ is the dual gauge field strength as defined in Eq. (2.5) in Section 2.2 of Chapter 2.

The Abelian gauge superfield can be generalized to the non-Abelian case. The vector superfield V then belongs to the adjoint representation of the gauge group, and is invariant under the supersymmetric generalized non-Abelian gauge transformation given by:

$$e^{-2gV'} = e^{2ig\Lambda^\dagger} e^{-2gV} e^{-2ig\Lambda}, \tag{3.57}$$

where Λ is defined in Eq. (3.44), g is the gauge coupling, and the vector superfield V is now represented as a matrix:

$$V_{ij} = T^a_{ij} V_a, \tag{3.58}$$

where T^a are the Hermitian generators of the gauge group. Selecting the Wess–Zumino gauge here preserves only the ordinary non-Abelian gauge symmetry contained within the generalized supersymmetry gauge transformation Eq. (3.57). The Wess–Zumino gauge does not fix the non-Abelian gauge symmetry. The supersymmetric field strength for V

can also be generalized to non-Abelian transformations, with the result that W_α, now the non-Abelian gauge superfield strength, obeys:

$$W_\alpha = -\frac{1}{8g}\bar{D}^2 e^{2gV} D_\alpha e^{-2gV}, \tag{3.59}$$

$$W_\alpha \rightarrow W'_\alpha = e^{2ig\Lambda} W_\alpha e^{-2ig\Lambda}, \tag{3.60}$$

$$\bar{W}_{\dot{\alpha}} \rightarrow \bar{W}'_{\dot{\alpha}} = e^{2ig\Lambda^\dagger} \bar{W}_{\dot{\alpha}} e^{-2ig\Lambda^\dagger}. \tag{3.61}$$

The component form of W is:

$$W_\alpha = T^a \left(-i\lambda^a_\alpha + \theta_\alpha D^a - \frac{i}{2}(\sigma^m \bar{\sigma}^n \theta)_\alpha v^a_{mn} + \theta^2 \sigma^m D_m \bar{\lambda}^a\right), \tag{3.62}$$

where v^a_{mn} is the non-Abelian gauge field strength, and D_m is the non-Abelian covariant derivative, respectively given by and acting as:

$$v^a_{mn} = \partial_m A^a_n - \partial_n A^a_m + igf^{abc} A^b_m A^c_n, \quad D_m \bar{\lambda}^a = \partial_m \bar{\lambda}^a + gf^{abc} A^b_m \bar{\lambda}^c. \tag{3.63}$$

Given chiral and vector superfields Φ and V and the vector superfield strength W_α, the Lagrangians defining classical renormalizable supersymmetric gauge theories can be constructed in their entirety. In Section 3.3 we will utilize chiral and vector superfields to construct supersymmetric gauge theories.

3.3 $\mathcal{N} = 1$ Supersymmetric Gauge Theories

In Subsection 3.2.3 of Section 3.2, the superfield formalism was used to construct $\mathcal{N} = 1$ supersymmetry covariant fields with specific properties. From these fields, the most general supersymmetric renormalizable Lagrangians for these separate fields can be constructed [113, 114, 117]. Supersymmetric Lagrangians which specify supersymmetric field theories can be obtained as the highest order component in the variables $(\theta, \bar{\theta})$ of a superfield. The θ^2 and $\bar{\theta}^2$ components of a chiral superfield (or product of chiral superfields, also a chiral superfield) transform via a supersymmetry transformation into a total derivative. The result of such integrations gives zero by the total derivative vanishing at the boundaries of integration. An integrand which exhibits this behaviour is then invariant under supersymmetry.

We shall now describe the general Lagrangians for chiral and vector multiplets separately. When these are coupled together in a gauge invariant way, the result will be the Lagrangian for $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory. Like other supersymmetric gauge theories, $\mathcal{N} = 1$ supersymmetric gauge theories has a vanishing vacuum energy, which is brought about by supersymmetry.

Given an $\mathcal{N} = 1$ supersymmetric scalar multiplet, the most general Lagrangian including interactions is given by:

$$\mathcal{L}_{\text{scalar}} = \int d^4\theta K(\Phi, \Phi^\dagger) + \int d^2\theta \mathcal{W}(\Phi) + \int d^2\bar{\theta} \bar{\mathcal{W}}(\Phi^\dagger), \tag{3.64}$$

where the non-holomorphic function $K(\Phi, \Phi^\dagger)$ is known as the Kähler potential, and $\mathcal{W}$ is the superpotential, as introduced in Eq. (3.38) of Subsection 3.2.3. The Kähler potential determines the metric on the field space, g^{ij}, defined by $g^{ij} = \partial^2 K/\partial A_i \partial A_j$. As noted previously, the auxiliary fields F_i which appear in $\mathcal{L}_{\text{scalar}}$ can be eliminated algebraically through the Euler-Lagrange equations.

To ensure that $\mathcal{L}_{\text{scalar}}$ is a renormalizable Lagrangian, the functions K and $\mathcal{W}$ are constrained by R-symmetry, as defined in Eqs. (3.25,3.26) of Subsection 3.2.3. However, we note that R-symmetry is not a necessary condition for renormalizability, but it is a sufficient one. The R-symmetry acts on chiral superfields as follows:

$$R\Phi(x,\theta) = \Phi'(x,\theta) = e^{2in\alpha}\Phi(x, e^{-i\alpha}\theta), \tag{3.65}$$

$$R\Phi^\dagger(x,\bar{\theta}) = \Phi'^\dagger(x,\bar{\theta}) = e^{-2in\alpha}\Phi^\dagger(x, e^{i\alpha}\bar{\theta}). \tag{3.66}$$

The component fields of the chiral superfields transform under the R-symmetry according to:

$$\phi \rightarrow e^{2in\alpha}\phi, \tag{3.67}$$

$$\psi \rightarrow e^{2i(n-1/2)\alpha}\psi, \tag{3.68}$$

$$F \rightarrow e^{2i(n-1)\alpha}F, \tag{3.69}$$

where the number n is the R-character of the R-symmetry. As the R-symmetry acts as $\theta \rightarrow e^{i\alpha}\theta$ and $d^2\theta \rightarrow e^{-2i\alpha}d^2\theta$, the R-character of superfields in $\mathcal{W}$ must sum to unity, and in K the R-characters must sum to zero (K is said to be 'R-neutral').

For the vector multiplet, the product of the Abelian vector superfield strengths $W_\alpha W^\alpha$ has a highest order term which can be expressed in component fields as:

$$W^\alpha W_\alpha|_{\theta\theta} = -2i\lambda\sigma^m\partial_m\bar{\lambda} + D^2 - \frac{1}{2}v^{mn}v_{mn} + \frac{i}{2}{}^*v_{kl}v_{kl}, \tag{3.70}$$

where ${}^*v_{kl}$ is the dual gauge field strength as defined in Section 2.2 of Chapter 2. Using this product of superfields, the Abelian supersymmetric vector field Lagrangian is given by:

$$\mathcal{L}_{\text{vector}} = \frac{1}{4g^2}\left(\int d^2\theta W^\alpha W_\alpha + \int d^2\bar{\theta}\bar{W}_{\dot{\alpha}}\bar{W}^{\dot{\alpha}}\right), \tag{3.71}$$

This Lagrangian can be generalized to the non-Abelian case by generalizing the product of Abelian field strengths $W_\alpha W^\alpha$ to:

$$\mathrm{Tr}(W^\alpha W_\alpha|_{\theta\theta}) = -2i\lambda^a\sigma^m D_m\bar{\lambda}^a + D^aD^a - \frac{1}{2}v^{amn}v^a_{mn} + \frac{i}{2}{}^*v^{akl}v^a_{kl}. \tag{3.72}$$

The non-Abelian supersymmetric vector field Lagrangian is then given by:

$$\mathcal{L}_{\text{vector}} = \frac{1}{4g^2}\mathrm{Tr}\left(\int d^2\theta W^\alpha W_\alpha + \int d^2\bar{\vartheta}\bar{W}_{\dot{\alpha}}\bar{W}^{\dot{\alpha}}\right). \tag{3.73}$$

The most general vector multiplet Lagrangian will contain a ϑ-term, defined in Chapter 2, given by $i\vartheta k$, where ϑ is the vacuum angle. The ϑ-term appears in the complexified gauge coupling τ, which has the same definition given previously in Eq. (2.17) of Section 2.2 in Chapter 2, where:

$$\tau = \frac{4\pi i}{g^2} + \frac{\vartheta}{2\pi}. \tag{3.74}$$

The addition of a ϑ-term to the Lagrangian Eq. (3.73) means that the non-Abelian vector multiplet Lagrangian can be written in the following compact form:

$$\begin{aligned}\mathcal{L}_{\text{vector}} &= \frac{1}{8\pi}\mathrm{Im}\left(\tau\mathrm{Tr}\int d^2\theta W^\alpha W_\alpha\right)\\ &= -\frac{1}{4g^2}v^a_{mn}v^{amn} + \frac{\vartheta}{32\pi^2}v^a_{mn}{}^*v^{amn} + \frac{1}{2g^2}(D^aD^a - 2i\lambda^a\sigma^m D_m\bar{\lambda}^a),\end{aligned} \tag{3.75}$$

in which τ, the complexified coupling, can be regarded as a chiral superfield.

In addition to the chiral and vector superfields describing pure Yang–Mills supersymmetric gauge theory (that is, pure non-Abelian supersymmetric gauge theory), one can add matter to the supersymmetric Yang–Mills gauge theories under consideration. Matter in these theories can be added in the form of chiral multiplets. Matter fields generally transform in the fundamental and anti-fundamental (or conjugate-fundamental) representation. Thus chiral superfields in the fundamental representation must be introduced if they are to form the matter content of the supersymmetric gauge theories under consideration. The chiral superfield Φ transforms in the adjoint representation, as given in Eq. (3.43) of Subsection 3.2.3.

We denote chiral superfields residing in the fundamental representation and anti-fundamental representation as Q and $\bar{Q}$, respectively. The component field content of Q and $\bar{Q}$ are denoted (q, χ, G) and $(\tilde{q}, \tilde{\chi}, \tilde{G})$, respectively. The component field content for the adjoint chiral superfields Φ and $\Phi^\dagger$ are respectively (ϕ, ψ, F) and $(\phi^\dagger, \bar{\psi}, F^\dagger)$, as given in Eqs. (3.39,3.40) of Subsection 3.2.3. The chiral matter superfield Q is a column vector of dimension N when the fundamental representation is N-dimensional. The superfield $\bar{Q}$ is an N-dimensional row vector in the dimension N fundamental representation. These chiral superfields obey the generalized supersymmetry gauge transformations:

$$Q \to Q' = e^{2ig\Lambda}Q, \quad Q^\dagger \to Q^{\dagger\prime} = Q^\dagger e^{-2ig\Lambda^\dagger}, \tag{3.76}$$

$$\tilde{Q} \to \tilde{Q}' = \tilde{Q}e^{-2ig\Lambda}, \quad \tilde{Q}^\dagger \to \tilde{Q}^{\dagger\prime} = e^{2ig\Lambda^\dagger}\tilde{Q}^\dagger, \tag{3.77}$$

where g is the gauge coupling.

In the Lagrangians which follow, we suppress the gauge group ('colour') indices for clarity.

$\mathcal{N} = 1$ Supersymmetric Yang–Mills theory

We now give the Lagrangian which specifies $\mathcal{N} = 1$ supersymmetric Yang–Mills theory, which is based upon the $\mathcal{N} = 1$ vector multiplet [113, 117] described in Subsection 3.2.3. The complete $\mathcal{N} = 1$ supersymmetric Yang–Mills theory Lagrangian is given by:

$$\mathcal{L}_{\mathcal{N}=1\ \mathrm{SYM}} = \frac{1}{8\pi}\mathrm{Im}\left(\tau\mathrm{Tr}\int d\theta W^\alpha W_\alpha\right) + \int d^2\theta d^2\bar{\theta}\,\Phi^\dagger e^{-2gV}\Phi + \int d^2\theta\mathcal{W} + \int d^2\bar{\theta}\bar{\mathcal{W}}. \tag{3.78}$$

The term $\Phi^\dagger e^{-2gV}\Phi$ in the Lagrangian $\mathcal{L}_{\mathcal{N}=\infty\ \mathrm{SYM}}$ is the gauge invariant kinetic terms for the chiral superfields. The relative normalization between these terms and the other terms in the $\mathcal{L}_{\mathcal{N}=1\ \mathrm{SYM}}$ is not fixed by $\mathcal{N} = 1$ supersymmetry since each term is by itself $\mathcal{N} = 1$ covariant. The proper relative normalization is obtained when all fermionic kinetic

terms in the Lagrangian have the same coefficients. In the case above, the normalization of the scalar kinetic term has been set to unity. We note that the complexified gauge coupling τ can be interpreted as a chiral superfield in the Lagrangian in Eq. (3.78).
The Lagrangian $\mathcal{L}_{\mathcal{N}=1\ \mathrm{SYM}}$ defined by Eq. (3.78) is invariant under the generalized supersymmetric gauge transformation Eq. (3.57) of Subsection 3.2.3.

$\mathcal{N}=1$ Supersymmetric QCD

One can generalize the Lagrangian of $\mathcal{N}=1$ supersymmetric Yang–Mills theory by including chiral matter multiplets. The chiral multiplet and the vector multiplet must couple to each other and remain invariant under the generalized supersymmetry gauge transformation. The resulting quantum version of the theory is referred to as $\mathcal{N}=1$ supersymmetric quantum chromodynamics, or $\mathcal{N}=1$ supersymmetric QCD (SQCD) since it is the $\mathcal{N}=1$ supersymmetric generalization of quantum chromodynamics (QCD).
The Lagrangian of $\mathcal{N}=1$ SQCD is given by:

$$\mathcal{L}_{\mathcal{N}=1\ \mathrm{SQCD}} = \mathcal{L}_{\mathcal{N}=1\ \mathrm{SYM}} + \mathcal{L}_{\mathrm{matter}}, \tag{3.79}$$

where $\mathcal{L}_{\mathcal{N}=1\ \mathrm{SYM}}$ is given by Eq. (3.78) and $\mathcal{L}_{\mathrm{matter}}$ is the matter Lagrangian composed of the fundamental chiral superfields Q and $\tilde{Q}$ which transform according to Eqs. (3.76,3.77).
A sufficiently general structure for $\mathcal{L}_{\mathrm{matter}}$ is one which couples N_f sets or 'flavours' of fundamental chiral multiplets Q_f and N_f flavours of anti-fundamental chiral multiplets $\tilde{Q}_f$ via a mass term, which imparts mass to both Q_f and $\tilde{Q}_f$:

$$\begin{aligned}\mathcal{L}_{\mathrm{matter}} = &\int d^2\theta d^2\bar{\theta} \sum_{f=1}^{N_f} \left(Q_f^\dagger e^{-2gV} Q_f + \tilde{Q}_f e^{2gV} \tilde{Q}_f^\dagger \right) \\ &+ \left[\int d^2\theta \sum_{f=1}^{N_f} m_f \tilde{Q} Q_f + \int d^2\bar{\theta} \sum_{f=1}^{N_f} m_f \tilde{Q}^\dagger Q_f^\dagger \right],\end{aligned} \tag{3.80}$$

where f runs from 1 to N_f and m_f is the mass of the f^{th} chiral matter multiplet.
In the action for the matter Lagrangian $\mathcal{L}_{\mathrm{matter}}$, the component fields $(G, \tilde{G}, F)$ do not contribute to the dynamics of the theory and can be integrated out of the Lagrangian $\mathcal{L}_{\mathcal{N}=1\ \mathrm{SQCD}}$; they are thus auxiliary fields.

3.4 $\mathcal{N} = 2$ Supersymmetric Gauge Theories

In past theoretical developments, the motivation for considering supersymmetry with $\mathcal{N} > 1$ generators was the need to look at non-Abelian gauged supergravity models in the pursuit of unifying the known physical forces, including gravity, through supersymmetry. However, that $\mathcal{N} = 2$ supersymmetry is the first example of extended supersymmetry is also important. In Chapter 5 we shall describe exact results in $\mathcal{N} = 2$ supersymmetric gauge theories. In this section we outline the construction of $\mathcal{N} = 2$ supersymmetric gauge theories in the same way in which $\mathcal{N} = 1$ supersymmetry was described in Section 3.3.

Classical Yang–Mills gauge theory with $\mathcal{N} = 2$ extended supersymmetry can be constructed within the $\mathcal{N} = 1$ supersymmetry formalism or by starting from the $\mathcal{N} = 2$ supersymmetry formalism. We begin with the former as it is closer to the theory described in Section 3.3, and then proceed to formulate $\mathcal{N} = 2$ supersymmetric Yang–Mills theory in terms of $\mathcal{N} = 2$ supersymmetry.

The field multiplets described in Section 3.3, namely the on-shell $\mathcal{N} = 1$ scalar multiplet and the $\mathcal{N} = 1$ vector multiplet, have the same field content as the on-shell $\mathcal{N} = 2$ vector multiplet. Excluding auxiliary fields, the $\mathcal{N} = 2$ supersymmetric field strength Ψ contains all of the fields present together in the $\mathcal{N} = 1$ gauge invariant superfields Φ and W_α, that is, the field multiplet (A, ψ, λ, v_m). The most general $\mathcal{N} = 1$ supersymmetric Yang–Mills Lagrangian does contain all of these component fields, but it is not $\mathcal{N} = 2$ supersymmetric. Further restrictions arise when we demand that $\mathcal{L}_{\mathcal{N}=1\ \mathrm{SYM}}$ possesses $\mathcal{N} = 2$ supersymmetry. The conditions on the Lagrangian $\mathcal{L}_{\mathcal{N}=1\ \mathrm{SYM}}$ which ensure that it now has $\mathcal{N} = 2$ supersymmetry are: (i) that the fields A_i and ψ_i belong to the same gauge group representation as v_m^a and λ^a, (that is, Φ must transform in the adjoint representation of the gauge group) and (ii) that, for the case when there are no matter multiplets present, the $\mathcal{N} = 1$ superpotential $\mathcal{W}$ is set to zero. The second condition sets the previously arbitrary relative normalization between the Yang–Mills term and the scalar term in the Lagrangian to be equal: this can be done by rescaling the chiral superfield by the coupling as $\Phi \to \Phi/g$. Matter in the form of fundamental or anti-fundamental chiral multiplets Q_f and $\tilde{Q}_f$ can be included in the theory by coupling the pure $\mathcal{N}=2$ supersymmetric Yang–Mills gauge theory in an identical fashion to that done for the $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory in Section 3.3. Historically, $\mathcal{N}=2$ matter multiplets were known as matter hypermultiplets, to distinguish them from $\mathcal{N}=1$ multiplets. We shall refer to the hypermultiplets as $\mathcal{N}=2$ matter multiplets in keeping with Section 3.3. In this section we refer to the original papers [112, 118] and also the reviews [121, 122, 123].

$\mathcal{N}=2$ Supersymmetric Yang–Mills theory

The $\mathcal{N}=2$ supersymmetric Yang–Mills Lagrangian in terms of $\mathcal{N}=1$ superspace is:

$$\mathcal{L}_{\mathcal{N}=2\ \mathrm{SYM}} = \frac{1}{8\pi}\mathrm{Im}\left(\tau\mathrm{Tr}\left[\int d^2\theta\, W^\alpha W_\alpha + 2\int d^2\theta d^2\bar{\theta}\,\Phi^\dagger e^{2gV}\Phi e^{-2gV}\right]\right). \tag{3.81}$$

Those terms in $\mathcal{L}_{\mathcal{N}=2\ \mathrm{SYM}}$ involving auxiliary fields can be collected to give the expression:

$$\frac{1}{g}\mathrm{Tr}\left(\frac{1}{2}DD + D[\phi^\dagger,\phi] + F^\dagger F\right),$$

wherein each field exists in the adjoint representation of the gauge group. The auxiliary fields D and F can be eliminated to give the scalar potential dependent upon the scalar field ϕ identified with a Higgs field:

$$V(\phi) = -\frac{1}{2g^2}\mathrm{Tr}\big([\phi^\dagger,\phi]^2\big). \tag{3.83}$$

The Lagrangian $\mathcal{L}_{\mathcal{N}=2\ \mathrm{SYM}}$ can be expanded in terms of component fields as:

$$\begin{aligned}\mathcal{L}_{\mathcal{N}=2\ \mathrm{SYM}} = \frac{1}{g^2}\mathrm{Tr}\bigg(&-\frac{1}{4}v_{mn}v^{mn} + g^2\frac{\theta}{32\pi^2}v_{mn}{}^*v^{mn} + (D_m\phi)^\dagger D^m\phi - i\sqrt{2}[\lambda,\psi]\phi^\dagger \\ &-i\sqrt{2}[\bar{\lambda},\bar{\psi}]\phi - \frac{1}{2}[\phi^\dagger,\phi]^2 - i\lambda\sigma^m D_m\bar{\lambda} - i\bar{\psi}\bar{\sigma}^m D_m\psi\bigg).\end{aligned} \tag{3.84}$$

in which the auxiliary fields D and F do not appear, having been eliminated.

To formulate $\mathcal{N}=2$ supersymmetric Yang–Mills gauge theory in $\mathcal{N}=2$ superspace, one extends the previous $\mathcal{N}=1$ superspace to $\mathcal{N}=2$ superspace by introducing four more Grassmannian degrees of freedom to the set of $\mathcal{N}=1$ supersymmetry internal degrees of freedom. Let the $\mathcal{N}=2$ superspace be the set of co-ordinates $\{x, \theta, \bar{\theta}, \tilde{\theta}, \bar{\tilde{\theta}}\}$. Then a general $\mathcal{N}=2$ superfield is the function $F(x, \theta, \bar{\theta}, \tilde{\theta}, \bar{\tilde{\theta}})$. To obtain the same component field content as the $\mathcal{N}=2$ vector multiplet, the constraints of chirality and reality can be imposed on a general $\mathcal{N}=2$ superfield. Let the resulting $\mathcal{N}=2$ vector superfield be Ψ. This then permits one to express the Lagrangian $\mathcal{L}_{\mathcal{N}=2 \text{ SYM}}$ as:

$$\mathcal{L}_{\mathcal{N}=2 \text{ SYM}} = \frac{1}{4\pi} \text{Im}\left(\text{Tr} \int d^2\theta d^2\tilde{\theta} \, \frac{1}{2} \tau \Psi^2 \right), \tag{3.85}$$

where Ψ is the $\mathcal{N}=2$ vector superfield. The most general Lagrangian for $\mathcal{N}=2$ supersymmetric Yang–Mills theory can be written in terms of a function $\mathcal{F}(\Psi)$, known as the $\mathcal{N}=2$ prepotential. The Lagrangian $\mathcal{L}_{\mathcal{N}=2 \text{ SYM}}$ then reads:

$$\begin{aligned}\mathcal{L}_{\mathcal{N}=2 \text{ SYM}} &= \frac{1}{4} \text{Im}\left(\text{Tr} \int d^2\theta d^2\tilde{\theta} \, \mathcal{F}(\Psi) \right) \\ &= \frac{1}{8\pi} \text{Im}\left(\int d^2\theta \, \frac{\partial^2 \mathcal{F}(\Phi)}{\partial \Phi^a \partial \Phi^b} W^{a\alpha} W^b_\alpha + 2 \int d^2\theta d^2\tilde{\theta} \, (\Phi^\dagger e^{2gV})^a \frac{\mathcal{F}(\Phi)}{\partial \Phi^a} \right). \end{aligned} \tag{3.86}$$

The Kähler potential for $\mathcal{N}=2$ supersymmetric Yang–Mills theory can be read off from Eq. (3.86) as $\text{Im}(\Phi^{\dagger a} \mathcal{F}_a(\Phi))$. Then the metric on the space of fields is given by $g_{ab} = \text{Im}(\partial_a \partial_b \mathcal{F})$; metrics of this form are known as special Kähler metrics. Renormalizablity of the theory is ensured by demanding that the prepotential $\mathcal{F}$ is quadratic in the $\mathcal{N}=2$ vector superfield Φ. However, for an effective $\mathcal{N}=2$ supersymmetric Yang–Mills theory, this condition need not hold in order that the theory be renormalizable.

$\mathcal{N}=2$ Supersymmetric QCD

Pure $\mathcal{N}=2$ supersymmetric Yang–Mills gauge theory coupled to chiral $\mathcal{N}=2$ matter multiplets results in $\mathcal{N}=2$ supersymmetric quantum chromodynamics, or $\mathcal{N}=2$ supersymmetric QCD (SQCD). In terms of $\mathcal{N}=1$ superspace, the general $\mathcal{N}=2$ supersymmetric QCD Lagrangian is given by:

$$\mathcal{L}_{\mathcal{N}=2 \text{ SQCD}} = \mathcal{L}_{\mathcal{N}=2 \text{ SYM}} + \mathcal{L}_{\text{matter}} + \left[\sqrt{2} i g \int d^2\theta \sum_{f=1}^{N_f} \tilde{Q}_f \Phi Q_f - \sqrt{2} i g \int d^2\bar{\theta} \sum_{f=1}^{N_f} Q_f^\dagger \Phi^\dagger \tilde{Q}_f \right], \tag{3.87}$$

where the Lagrangian $\mathcal{L}_{\mathcal{N}=2 \text{ SYM}}$ is given by Eq. (3.81), and $\mathcal{L}_{\text{matter}}$ is given by Eq. (3.80) of Section 3.3. The term additional to the Lagrangians $\mathcal{L}_{\mathcal{N}=2 \text{ SYM}}$ and $\mathcal{L}_{\text{matter}}$ within

$\mathcal{L}_{\mathcal{N}=2\ \mathrm{SQCD}}$ can be intepreted as the supersymmetric generalization of a Yukawa coupling term.

Exact results proposed for $\mathcal{N}=2$ supersymmetric Yang–Mills theory and $\mathcal{N}=2$ supersymmetric QCD shall be described in Chapter 5.

3.5 $\mathcal{N}=3$ Supersymmetric Gauge Theories

In this section we briefly describe $\mathcal{N}=3$ supersymmetric gauge theory [133]. The defining characteristic of these theories is that the supersymmetry constraints described in Subsection 3.2.2 for $\mathcal{N}=3$ supersymmetry reduce to the supersymmetric equations of motion and are entirely equivalent to them [133, 134].

Gauge theories with $\mathcal{N}=3$ supersymmetry are ultra-violet finite and scale invariant theories with vanishing beta function. The supersymmetry constraints for $\mathcal{N}=3$ supersymmetric gauge theories have the same form as the $\mathcal{N}=2$ supersymmetry as given in Eqs. (3.15) of Subsection 3.2.2, namely [119]:

$$\begin{aligned} v^{ij}_{\alpha\beta}+v^{ji}_{\beta\alpha} &= 0, \\ \bar{v}_{\dot{\alpha}i\dot{\beta}j}+\bar{v}_{\dot{\beta}i\dot{\alpha}j} &= 0, \\ v^{i}_{\alpha\dot{\beta}j} &= 0, \end{aligned} \tag{3.88}$$

where v^{ij}, $\bar{v}_{\dot{\alpha}i\beta j}$ and $v^{i}_{\alpha\dot{\beta}j}$ are Yang–Mills superfield strengths, with $i,j=1,2,3$.

An off-shell formulation of $\mathcal{N}=3$ supersymmetric Yang–Mills gauge theory is possible and requires an infinite set of auxiliary fields. The corresponding on-shell formulation of the theory is not known. To express the Lagrangian for off-shell $\mathcal{N}=3$ supersymmetric Yang–Mills gauge theory requires the use of harmonic $\mathcal{N}=3$ superspace [133], the details and construction of which we do not include here (for reviews see [134]).

There exist three harmonic $\mathcal{N}=3$ supersymmetric Yang–Mills spinor potentials or superconnections for $\mathcal{N}=3$ supersymmetric Yang–Mills gauge theory, which following Subsection 3.2.2 we denote as $\mathcal{A}^{i}_{\alpha}$, where $i=1,2,3$ and α is a Weyl index. These harmonic superconnections obey the conditions:

$$\mathcal{A}^{1\dagger}_{\alpha} = -\mathcal{A}^{1}_{\alpha}, \quad \mathcal{A}^{2\dagger}_{\alpha} = \mathcal{A}^{3}_{\alpha}, \tag{3.89}$$

and can be expanded in terms of component fields, of which there are an infinite number. After the Wess–Zumino gauge has been selected in this theory, there remain an infinite number of auxiliary fields contained within the $\mathcal{N} = 3$ harmonic superconnections. This excess of auxiliary degrees of freedom makes the theory unphysical and unsuitable as a candidate for extending known theories. There also exist three harmonic $\mathcal{N} = 3$ supersymmetry covariant derivatives associated with each of the $\mathcal{N} = 3$ superconnections. In the notation of Subsection 3.2.2, the $\mathcal{N} = 3$ supersymmetry covariant derivatives will be the harmonic versions of the covariant derivatives $\mathcal{D}^i_\alpha$ and $\bar{\mathcal{D}}^i_{\dot{\alpha}}$ defined in Eqs. (3.10,3.11). Here we use this notation to denote the three harmonic $\mathcal{N} = 3$ supersymmetry covariant derivatives, with the same index range for i.

The Lagrangian for off-shell $\mathcal{N} = 3$ supersymmetric Yang–Mills gauge theory can then be expressed as:

$$\begin{aligned} \mathcal{L}_{\mathcal{N}=3\ \mathrm{SYM}} &= \mathrm{Tr}\left(\mathcal{A}^3(\mathcal{D}^1\mathcal{A}^2 - \mathcal{D}^2\mathcal{A}^1) - \mathcal{A}^2(\mathcal{D}^1\mathcal{A}^3 - \mathcal{D}^3\mathcal{A}^1)\right. \\ &\quad \left. + \mathcal{A}^1(\mathcal{D}^2\mathcal{A}^3 - \mathcal{D}^3\mathcal{A}^2) - (\mathcal{A}^1)^2 + 2i\mathcal{A}^1\left[\mathcal{A}^2, \mathcal{A}^3\right]\right), \end{aligned} \tag{3.90}$$

in which all indices except the supersymmetry ones have been suppressed. It can be shown that the Lagrangian $\mathcal{L}_{\mathcal{N}=3\ \mathrm{SYM}}$ contains the same on-shell dynamics which follow from the constraint formulation of the theory in Eq. (3.88).

Theories with $\mathcal{N} = 3$ supersymmetry are not useful theories due to the infinite number of auxiliary fields which they contain. Theories with $\mathcal{N} = 1, 2, 4$ supersymmetry are theories with phenomenological potential and are the main focus for the construction of supersymmetric models of elementary particles. We do not consider $\mathcal{N} = 3$ supersymmetric Yang–Mills gauge theory in later chapters of this thesis, but include it here for comparison and completeness.

3.6 $\mathcal{N} = 4$ Supersymmetric Gauge Theories

Theories with $\mathcal{N} = 4$ supersymmetry are referred to as having maximal supersymmetry. This is because supersymmetric field theories with $\mathcal{N} > 4$ supersymmetry do not possess asymptotic freedom and are not renormalizable. Furthermore, $\mathcal{N} > 4$ supersymmetry requires multiplets in which fields with spin $\frac{3}{2}$ or greater exist, whereas $\mathcal{N} = 4$ supersymmetry contains fields with at most spin 1. Since particles with spin greater than one have not yet been observed, this makes $\mathcal{N} = 4$ supersymmetric field theories the limiting case for most phenomenological applications of supersymmetry. Locally supersymmetric field theories with greater degrees of supersymmetry, and which thus contain higher spin particles, have also been used to model gravitation, such as $\mathcal{N} = 8$ supergravity theories. The sign of the beta function in supersymmetric field theories changes once the theory has more than $\mathcal{N} = 4$ degrees of supersymmetry. Theories with $\mathcal{N} = 4$ supersymmetry are ultra-violet finite: they are exactly scale-invariant theories with vanishing beta function. This makes $\mathcal{N} = 4$ supersymmetric gauge theories valuable from a theoretical perspective since they possess special properties.

Supersymmetric gauge theories have the property that differing degrees of supersymmetry can be studied within the same supersymmetric field theory by modifying the matter content or reducing the spacetime dimension of the theory. An example of the former was provided in Section 3.4, in which $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory was obtained from $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory with a particular matter content. In this subsection an example of the dimensional reduction method is provided as a method for obtaining $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory [123, 159, 191]. These methods for obtaining supersymmetric gauge theories from less supersymmetric ones permits one to circumvent the notational complexity which one could encounter in field theories with higher degrees of supersymmetry. In this section we follow the reviews of $\mathcal{N} = 4$ supersymmetric gauge theories in [123] and in [159, 191].

The $\mathcal{N} = 4$ superspace requires the introduction of four more Grassmann-valued parameters to be added to the four $\mathcal{N} = 2$ superspace co-ordinates already introduced. Given this requirement of $\mathcal{N} = 4$ supersymmetry, the method of dimensional reduction is useful as it simplifies the construction of $\mathcal{N} = 4$ supersymmetric gauge theories.

One can obtain four dimensional classical $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory by reducing the dimension of ten dimensional $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory from ten to four. (In a similar way, one can obtain four dimensional $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory by reducing the dimension of six dimensional

$\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory.) Ten dimensional classical $\mathcal{N} = 1$ supersymmetric Yang–Mills theory has the following Lagrangian:

$$\mathcal{L}_{D=10\ \mathcal{N}=1\ \text{SYM}} = -\frac{1}{4}v_{MN}v^{MN} - \frac{1}{2}\bar{\lambda}\Lambda^{M}(D_M\lambda), \tag{3.91}$$

where the gauge field A^a_M and spinor λ^a reside in the adjoint representation of the gauge group, and the ten dimensional Lorentz indices indices M, N run from 0 to 9. The following constraints on λ^a are also present:

$$(1-\Lambda_{11})\lambda = 0, \quad \bar{\lambda} = \lambda^T C_{10}, \tag{3.92}$$

where $\Lambda_{11} = \Lambda_0\Lambda_1\Lambda_2\cdots\Lambda_9$ and C_{10} is the charge conjugation operator in ten spacetime dimensions, such that $C_{10}\Lambda_M C_{10}^{-1} = -\Lambda^T_M$. The non-Abelian gauge field strength and covariant derivative are defined in the usual sense (but in ten spacetime dimensions) as:

$$v^a_{MN} = \partial_M v^a_N - \partial_N v^a_M + igf^{abc}v^b_M v^c_N, \quad D_M\lambda^a = \partial_m\lambda^a + gf^{abc}A^b_M\lambda^c. \tag{3.93}$$

To reduce the dimensions of the $\mathcal{N} = 1$ theory specified in Eq. (3.91), one demands that all of the fields present in the theory depend only on four dimensional spacetime. The ten dimensional spacetime co-ordinate x^M can be decomposed into the four dimensional spacetime co-ordinate x^m and a six-dimensional one, x^i, as $x^M = (x^m, x^i)$. This reduces the ten dimensional Lorentz group $SO(1,9)$ of the theory to the product $SO(1,3) \times SO(6)$. Consequently, the ten dimensional fields λ^a and v^a_m decompose into the set of four dimensional fields $\{\phi_{ij}, \lambda_i\}$. The result is that the Lagrangian of four dimensional classical $\mathcal{N} = 4$ supersymmetric Yang–Mills theory can be obtained from the dimensional reduction of the Lagrangian in Eq. (3.91), and can be written as:

$$\begin{aligned}\mathcal{L}_{\mathcal{N}=4\ \text{SYM}} = \operatorname{Tr}\Bigg(&-\frac{1}{4}v_{mn}v^{mn} + i\lambda_i\sigma^m D_m\bar{\lambda}_i + \frac{1}{2}D_m\phi_{ij}D^m\phi^{ij} \\ &+i\lambda_i[\lambda_j, \phi^{ij}] + i\bar{\lambda}^i[\bar{\lambda}_j, \phi_{ij}] + \frac{1}{4}[\phi_{ij}, \phi_{kl}][\phi^{ij}, \phi^{kl}]\Bigg).\end{aligned} \tag{3.94}$$

The second method by which $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory can be constructed involves coupling pure $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory to an $\mathcal{N} = 2$ matter multiplet which transforms in the adjoint representation. The resulting theory will have two sets of adjoint $\mathcal{N} = 2$ chiral multiplets and an $\mathcal{N} = 2$ vector multiplet. The field content of the modified $\mathcal{N} = 2$ theory will be identical to that of

$\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory, and thereby the Lagrangian $\mathcal{L}_{\mathcal{N}=4 \text{ SYM}}$ is obtained.

In Chapter 4, exact results in $\mathcal{N} = 4$ supersymmetric Yang–Mills theory will be described, including the conjectural Olive-Montonen electric-magnetic duality of the theory.

Chapter 4

Exact Results in Supersymmetric Gauge Theories I: $\mathcal{N} = 1$ and $\mathcal{N} = 4$ Supersymmetry

4.1 Introduction

In Chapter 3 the concept of global supersymmetry and the construction of supersymmetric gauge theories were described. The minimally and maximally supersymmetric generalizations of Yang–Mills gauge theory, respectively $\mathcal{N} = 1$ and $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory, were constructed. In this chapter we endeavour to briefly review some of the exact results obtained in these gauge theories, which the presence of supersymmetry makes possible. In doing so, we introduce some of the concepts which will be further described in Chapter 5 for the case of $\mathcal{N} = 2$ supersymmetric gauge theories. However, there is a greater variety of phenomenon in $\mathcal{N} = 1$ supersymmetric gauge theories than in $\mathcal{N} = 2$ theories. Consequently, our review of exact results in $\mathcal{N} = 1$ theories is brief and incomplete. In contrast, $\mathcal{N} = 4$ supersymmetric gauge theories are simpler than $\mathcal{N} = 2$ theories. Thus $\mathcal{N} = 2$ theories are, in terms of relative complexity, intermediate theories. We describe exact results in $\mathcal{N} = 2$ theories in Chapter 5.

In this chapter we describe some of the exact results and proposed exact results obtained in $\mathcal{N} = 1$ and $\mathcal{N} = 4$ non-Abelian gauge theories. In Section 4.2 the significant early and recent exact results in $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory are outlined. These include non-renormalization theorems and the exact calculation of the beta function and various condensates. Section 4.3 briefly describes the exact results obtained in $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory. In Section 4.4 the phenomenon of electric-magnetic duality in $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory is described. This duality is conjectural, but there exists noteworthy evidence for its existence. In Section 4.5 we describe a special form of duality which occurs in $\mathcal{N} = 1$ supersymmetric gauge theories. This duality is known as Seiberg duality, after the author who uncovered this phenomenon. Although Seiberg duality is not associated with electric-magnetic duality, which shall reappear again in Chapter 5 in the context of $\mathcal{N} = 2$ supersymmetric theories, we include it here as part of our review concerning duality in supersymmetric gauge theories.

We continue to use the conventions and notation of Chapter 3. Useful reviews of the exact results in $\mathcal{N} = 1$ supersymmetric gauge theories include [123, 138, 139, 140]. Works which include reviews of results in $\mathcal{N} = 4$ supersymmetric gauge theories include [191, 159]. Upon a first reading, the reader may omit this chapter, and in particular, Section 4.4 and Section 4.5, without loss of essential material.

4.2 Exact Results in $\mathcal{N} = 1$ Supersymmetric Gauge Theories

The single supersymmetry of $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory permits one to exactly calculate various physically relevant quantities in the theory. In quantum $\mathcal{N} = 1$ supersymmetric gauge theory, this fact is important since very few results in any four dimensional quantum field theory have been determined exactly. In general, only with the simplifying constraint of supersymmetry have quantum field theories which bear some resemblance to phenomenologically relevant quantum field theories, such as quantum chromodynamics, yielded any exactly calculable quantities. These exact results are the first such results ever found in dynamically non-trivial four dimensional quantum field theories.

In this section we begin by briefly describing the earliest exact results in $\mathcal{N} = 1$ supersymmetric gauge theories. These are particular dynamical quantities and observables of these theories which were obtained exactly by a variety of methods. The gluino condensate and the beta function in the theory are historically the first examples of non-trivial quantities exactly calculated in the strong coupling regime of a four dimensional quantum field theory (although in the case of the gluino condensate the calculation was actually erroneous). In this section we refer to the original papers in [135, 136, 137, 143, 142], the later works in [141, 144], and the reviews in [138, 139, 140].

We now describe various exact results and properties of $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory, some of which are valid or extendable to $\mathcal{N} = 1$ theories with matter, such as $\mathcal{N} = 1$ SQCD. In some cases, the exact results hold only for the low energy effective theory. The effective action for these theories is the Wilsonian effective action rather than other effective actions such as the one-particle irreducible (1PI) effective action. The Wilsonian effective action at an energy scale μ is obtained by integrating out all of the fields appearing in the action which have mass greater than μ and the high momentum ($P > \mu$) modes of the 'light fields' (whose masses are less than μ). Unlike the 1PI effective action, the Wilsonian effective action has no infra-red ambiguities and no holomorphic anomalies [141]. The Wilsonian effective action must also possess the same global symmetries of the the microscopic action of the theory. In this thesis the only effective actions which we shall describe will be Wilsonian effective actions. One example of an exact result obtained in the low energy effective $\mathcal{N} = 1$ supersymmetric Yang–Mills theory is the form of the effective $\mathcal{N} = 1$ superpotential, denoted by $\mathcal{W}_{\text{eff}}$.

Vacuum Energy

The first exact result in any globally supersymmetric field theory, of which supersymmetric gauge theories are a subset, is the vanishing of the ground state energy or vacuum energy, E_{vac}. The vacuum energy is precisely zero order by order in the gauge coupling constant g, and receives no perturbative or non-perturbative quantum corrections:

$$E_{\text{vac}} = 0. \tag{4.1}$$

In this sense, the vanishing of the vacuum energy is an exact results, common to all supersymmetric gauge theories, and one which is required by the presence of supersymmetry [123].

The result in Eq. (4.1) arises due to the general expression for vacuum energy in supersymmetric field theories [136]:

$$E_{\text{vac}} \propto \int d^4x\, d^4\theta, \tag{4.2}$$

which is fixed by the uniformity of superspace. The vanishing of E_{vac} in Eq. (4.1) is ensured by the integration over $d^4\theta$ in the general expression above.

Non-renormalization Theorem

Another property of $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory which can be interpreted as an exact result is the $\mathcal{N} = 1$ supersymmetric non-renormalization theorem [135, 123, 140], which was also proven in [141]. Here we decribe the derivation first given in [135] and reviewed in [123, 140]. This particular derivation, first given in [135], uses perturbative techniques. In the latter derivation, the holomorphy of the $\mathcal{N} = 1$ superpotential is used, and the theorem is proven non-perturbatively [141].

For calculational and phenomenological purposes, one can derive the $\mathcal{N} = 1$ superfield Feynman rules from the action of $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory. The action of the theory will be the spacetime integral of the Lagrangian defining $\mathcal{N} = 1$ supersymmetric Yang–Mills gauge theory, given in Eq. (3.78) of Section 3.3 in Chapter 3. The $\mathcal{N} = 1$ superfield Feynman rules can be obtained by a superspace perturbation expansion of the action. This perturbation expansion will include terms of the form:

$$\int d^4x \int d^2\theta\, \Phi, \quad \int d^4x \int d^2\theta\, g\Phi^3, \tag{4.3}$$

and their conjugates. A problem arises for these terms because the adjoint chiral superfields Φ and $\Phi^\dagger$ are constrained scalar superfields, obeying the chirality conditions in Eqs. (3.30 ,3.36) in Subsection 3.2.3 of Chapter 3. Thus they do not exist across the entire $\mathcal{N} = 1$ superspace. Then the integrals in Eq. (4.3) can only be performed over part of the $\mathcal{N} = 1$ superspace, leaving the integration over the Grassmann parameters $(\theta, \bar{\theta})$ ill-defined.

For the case of the first integrals in Eq. (4.3), this problem can be overcome by introducing projection operators for the chiral superfields. These operators will project a chiral superfield out of a general scalar superfield, and so maps general scalar superfields to chiral superfields. The first form of integral in Eq. (4.3) can then be performed by converting the integration over part of the $\mathcal{N} = 1$ superspace to one over the entire $\mathcal{N} = 1$ superspace.

In the case of integrals of the second form in Eq. (4.3), the interation term $g\Phi^3$ will produce a superspace delta function which exists on half of the $\mathcal{N} = 1$ superspace. An identity can be used to write such delta functions over the full superspace, and integrals of the second form in Eq. (4.3) can be performed over the whole of the $\mathcal{N} = 1$ superspace. The superfield Feynman rules resulting from the perturbation expansion will have $\mathcal{N} = 1$ superspace dependence. From this, the form of arbitrary terms in the effective action for the $\mathcal{N} = 1$ theory can be deduced. These can be expressed as:

$$\int d^4\theta \int d^4x_1 \cdots d^4x_n \, F_1(x_1, \theta, \bar{\theta}) \cdots F_n(x_n, \theta, \bar{\theta}) G(x_1, \ldots, x_n), \tag{4.4}$$

where n is the number of external lines of an arbitrary one-particle irreducible Feynman supergraph, and F_n denote superfields and covariant derivatives of superfields. The function $G(x_1, \ldots, x_n)$ is a translationally invariant Green's function which contains all of the relevant spacetime structure.

The result in Eq. (4.4) is known as the $\mathcal{N} = 1$ non-renormalization theorem. This result holds for $\mathcal{N} = 1$ supersymmetric actions involving arbitrary numbers of vector and chiral superfields. An immediate consequence of Eq. (4.4) is that if external lines of a Feynman supergraph are chiral or anti-chiral, then the integral of Eq. (4.4) vanishes. Essentially, the $\mathcal{N} = 1$ non-renormalization theorem is an extension of the result of vanishing vacuum energy (Eq. (4.1)) to supersymmetric F-terms, which are the last components of chiral superfields. Since the $\mathcal{N} = 1$ superpotential $\mathcal{W}$ is composed only of F-terms, the superpotential of the theory is not renormalized at any order in the perturbation expansion. This is the statement which the name of the $\mathcal{N} = 1$ non-renormalization theorem refers to. However, the superpotential in general shall receive quantum non-perturbative corrections from instanton configurations.

A further consequence of the theorem is that all vacuum and tadpole Feynmann diagrams in the theory vanish. This is consistent with the exact vanishing of the vacuum energy E_{vac} in Eq. (4.1). In deriving the $\mathcal{N}=1$ non-renormalization theorem, one must regularize the spacetime loop integrals in a way consistent with supersymmetry, with an equal number of bosons and fermions in each case. This can be acheived with a regularization scheme known as dimensional reduction [127], which is particularly suitable for use in supersymmetric field theories.

NSVZ Beta Function and Wilsonian Beta Function

The beta function $\beta_{\mathcal{N}=1}(g)$ of quantum $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory can also be calculated exactly. The general beta function $\beta(g)$, as defined in Eq. (4.11) below, is the Gell-Mann–Low function of the theory and governs the running of the gauge coupling constant g. The Gell-Mann–Low function can be calculated to all orders in the coupling constant via a purely classical calculation (no quantum loop corrections are calculated). In the case of the low energy Wilsonian effective $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory, the beta function can be calculated exactly and is exact to one-loop quantum perturbative corrections.

One notable derivation of the exact Gell-Mann–Low function in quantum $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory is based on instanton calculus, and was first calculated by Novikov, Shifman, Vainshtein and Zakharov (NSVZ) for the gauge group $SU(N)$ [136]. These authors considered the one-instanton vacuum to vacuum transition in the $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory. At one-loop level in perturbation theory, the amplitude for the one-instanton transition is intrinsically supersymmetric. The supersymmetry of the classical $\mathcal{N}=1$ action is a symmetry of the theory which extends to the quantum $\mathcal{N}=1$ action and is present at all orders in the gauge coupling g. In addition there exists a non-renormalization theorem for instanton induced interactions in the theory which is directly analogous to the $\mathcal{N}=1$ non-renormalization theorem described above. This additional theorem asserts that the vacuum energy still vanishes in the presence of an external instanton field. These properties permit Novikov et al. to calculate the beta function exactly in quantum gauge theories with differing supersymmetry. Their result is known as the NSVZ beta function [136, 139].

The NSVZ beta function was derived for $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory with arbitrary classical gauge group and matter content. To extend the result to theories with extended supersymmetry, one can add additional $\mathcal{N}=1$ matter fields to form superfield multiplets with the same matter content as theories with extended supersymmetry [139]. Working within the Pauli–Villars regularization scheme [58], with arbitrary gauge group G and matter fields in an arbitrary gauge group representation $R=\sum_i R_i$ which is reducible, the most general NSVZ beta function has the form [139]:

$$\beta(g)_{\mathrm{NSZV}} = -\frac{g^2}{2\pi}\left[3T_G-\sum_i T(R_i)(1-\gamma_i)\right]\left(1-\frac{gT_G}{2\pi}\right)^{-1}, \tag{4.5}$$

where T_G and $T(R_i)$ are group factors and γ_i is the anomalous mass dimension of the matter fields. The objects T_G and $T(R_i)$ are defined as follows. Let T^a be the generators of the gauge group G in the representation R. Then one can define the factor $T(R)$ via $\mathrm{Tr}(T^aT^b)=T(R)\delta^{ab}$. The object $T(R)$ is known as the dual Coxeter number or one half of the Dynkin index for R. When R is the adjoint representation, we denote $T(R)=T_G$. When R is a reducible representation, one has $T(R)=T(R_i)$.

For gauge group $G=SU(N)$, we have that $T_G=N$. For matter fields transforming in the fundamental representation R_F of $SU(N)$, one has $T(R_F)=\frac{1}{2}$. Thus, for $\mathcal{N}=1$ supersymmetric $SU(N)$ Yang–Mills gauge theory with no matter multiplets, one has:

$$\beta_{\mathcal{N}=1}(g) = -\frac{3Ng^2}{2\pi}\left(1-\frac{gN}{2\pi}\right)^{-1}. \tag{4.6}$$

The beta function $\beta_{\mathcal{N}=1}(g)$ receives no perturbative or non-perturbative quantum corrections in the quantum $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory. Thus the beta function Eq. (4.6) is exact for quantum $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory. For the case of $\mathcal{N}=1$ supersymmetric QCD, the presence of matter multiplets modifies the NSVZ beta function as the effects of the the anomalous dimensions γ_i enter. The method used to derive the NSVZ beta function has been extended to the case of $\mathcal{N}=1$ supersymmetric QED in [137], which is an Abelian gauge theory. Instantons do not occur in Abelian (commutative) gauge theories, but the non-renormalization theorems derived by Novikov et al. [203, 202] can be extended to these theories.

The beta function for the low energy Wilsonian effective action of $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory can be obtained from the bare gauge coupling constant g as in Eq. (4.11) below. The Wilsonian beta function receives no quantum perturbative or non-perturbative corrections beyond the one loop level. In this sense, the Wilsonian beta function is one loop exact.

Gluino Condensate

The setting for these results is also pure $\mathcal{N} = 1$ supersymmetric $SU(N)$ quantum Yang–Mills theory. If the gauge field is identified with the gluon field in the theory, it is known as $\mathcal{N} = 1$ $SU(N)$ supersymmetric gluodynamics [138]. This is because the theory, although supersymmetric, resembles QCD with no quarks, and so is an isolated theory of gluons and their superpartners only. The Lagrangian for this theory has the same form as Eq. (3.79) in Section 3.3 with gauge group $SU(N)$. We now write the gauge field strength as $v^a_{mn} = G^a_{mn}$, where G^a_{mn} is the gluon field strength. The Lagrangian for $\mathcal{N} = 1$ supersymmetric gluodynamics is then:

$$\mathcal{L}_{\mathcal{N}=1\ SYM} = -\frac{1}{4g^2} G^a_{mn} G^{amn} + \frac{\vartheta}{32\pi^2} {}^{*}G^a_{mn} G^{amn} + \frac{i}{2g^2} \bar{\lambda}^a D_m \gamma^m \lambda^a. \qquad (4.7)$$

The superpartner of the gluon (vector) field in Eq. (4.7) is λ^a, known as the gluino field. Generically, the fermionic superpartner field of a gauge field is termed a gaugino field. Like QCD, $\mathcal{N} = 1$ supersymmetric gluodynamics is a strong coupling theory with asymptotic freedom and a confining phase. There exist Ward identities which constrain the theory. Some correlation functions can be deduced up to a constant factor. Examples of this include gauge invariant multi-point functions of the gluino operator in quantum $\mathcal{N} = 1$ supersymmetric gauge theory. The gluino field λ^a has a two-point correlator (a composite operator) which assumes a non-zero vacuum expectation value; this is the $\mathcal{N} = 1$ gluino condensate and was an early exact result.

The calculation of the $\mathcal{N} = 1$ gluino condensate has been the subject of some controversy [144]. This is because there once existed two apparently valid derivations of the condensate which were in disagreement. The first calculation of the condensate was performed using the semi-classical method; this is known as the strongly coupled instanton (SCI) approach. The second calculation uses the holomorphicity of the $\mathcal{N} = 1$ prepotential in the theory and is known as the weakly coupled instanton (WCI) approach.

The respective results from these approaches give the $\mathcal{N}=1$ $SU(N)$ gluino condensate as:

$$\frac{g^2}{16\pi^2}\langle \mathrm{Tr}\lambda^a\lambda_a\rangle_{\mathrm{SCI}} = \frac{2\Lambda^3}{[(N-1)!(3N-1)]^{1/N}}, \tag{4.8}$$

$$\frac{g^2}{16\pi^2}\langle \mathrm{Tr}\lambda^a\lambda_a\rangle_{\mathrm{WCI}} = \Lambda^3. \tag{4.9}$$

These two purportedly exact expressions for the gluino condensate are valid at both weak and strong coupling, and receive no perturbative or non-perturbative quantum corrections. However, they cannot both be correct because of this exactness.

In the strongly coupled instanton approach, the semi-classical method is employed. This method, which is strictly only valid at weak coupling, is used in the strong coupling regime of the $\mathcal{N}=1$ theory, which is a strongly coupled theory. The use of the semi-classical method in this regime is suspect since weak coupling is implicitly assumed within this method and although instanton effects become significant at strong coupling, their effects can only be calculated at weak coupling (without appeal to duality).

In comparison, the weakly coupled instanton approach extends the weakly coupled phase of the $\mathcal{N}=1$ theory to the strongly coupled phase by using the holomorphic property of the $\mathcal{N}=1$ prepotential $\mathcal{F}$. A weakly coupled method is not used and a valid calculation at strong coupling can be directly performed using the properties of the theory which arise from supersymmetry.

It has been shown that the result derived using the strongly coupled instanton method is erroneous. The $\mathcal{N}=1$ gluino condensate has now been calculated exactly for all simple gauge groups by Davies et al. [144]. This was achieved via one-monopole calculations in special settings, wherein monopoles assume the rôle of instantons.

Exact Superpotentials

That the $\mathcal{N}=1$ superpotential $\mathcal{W}$ is not renormalized perturbatively does not rule out the appearance of non-perturbative contributions to $\mathcal{W}$. From a phenomenological perspective, it is desirable that low energy supersymmetry has been broken non-perturbatively. The $\mathcal{N}=1$ non-renormalization theorem can be obtained via general arguments put forward by Seiberg [141, 146] and inspired by developments in string theory. These arguments can be used to extend perturbative non-renormalization theorems to non-perturbative cases, thereby providing information about the complete theory, which will include both perturbative and non-perturbative effects in general.

As described in Section 3.3 of Chapter 3, the complexified coupling constant τ can be interpreted as a chiral superfield in the $\mathcal{N}=1$ supersymmetric Yang–Mills Lagrangian. The coupling constants $\{g_i\}$ which appear in the superpotential $\mathcal{W}$ can also be considered as (background) chiral superfields. In the low energy effective theory, the effective superpotential $\mathcal{W}_{\text{eff}}$ is constrained by [141]:

1. The global symmetry group present when all the coupling constants are zero, $g_i = 0$.

2. Local holomorphicity of the superpotential: $\mathcal{W}_{\text{eff}}$ depends only on $\{g_i\}$ and not $\{g_i^\dagger\}$, treating the set of coupling constants $\{g_i\}$ as chiral superfields.

3. Asymptotic freedom of the couplings constants $\{g_i\}$ and the presence of a strongly coupled regime.

4. The weak coupling limit, in which $\{g_i\} \to 0$ for all i. Care must be taken in integrating out the chiral superfields $\{g_i\}$ when deriving the effective action, as $\mathcal{W}_{\text{eff}}$ may be non-analytic at the values $\{g_i\} = 0$.

These constraints are sufficient in many cases to derive non-renormalization theorems and some exact results, including the perturbative $\mathcal{N}=1$ non-renormalization theorem. Unusually, these lead to the conclusion that $\mathcal{W}_{\text{eff}}$ is in general a function not consistent with the global symmetries of the action from which the Wilsonian effective action is obtained.

An example of the use of the criteria given by Seiberg [141] is its application to the effective superpotential in the renormalized Wess–Zumino model, described briefly in Chapter 3. The effective superpotential $\mathcal{W}_{\text{eff}}$ can be shown, using the above constraints, to be equal to the tree-level superpotential of the theory. The superpotential receives no perturbative or non-perturbative quantum corrections to any order in the gauge coupling constant. Hence, in the Wess–Zumino model, the result $\mathcal{W}_{\text{eff}} = \mathcal{W}$, where $\mathcal{W}$ is the tree-level superpotential, is exact. Previously, the non-renormalization of the Wess–Zumino superpotential was known only perturbatively, which indicates that the above constraints can be used to obtain powerful results.

Given the $\mathcal{N}=1$ supersymmetric non-renormalization theorem, the objects in the theory which require renormalization are the chiral superfield Φ, the vector superfield V and the gauge coupling, which are renormalized respectively as follows:

$$\Phi_0^i \;=\; (Z^{1/2})^{ij}\Phi^j, \quad V_0 \;=\; Z_V^{1/2}V, \quad g_0 \;=\; Z_g g, \tag{4.10}$$

with the condition $Z_g Z_V^{1/2}=1$, where Z,Z_V and Z_g are the renormalization factors for the respective fields. The renormalized $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theory can be characterised by the beta function $\beta(g)$ and the anomalous dimensions matrix γ^{ij} of the chiral superfield Φ^i, which are defined as:

$$\beta(g) \;=\; \mu\frac{\partial g}{\partial \mu}, \tag{4.11}$$

$$\gamma^{ij} \;=\; \mu(Z^{-1/2})^{ik}\frac{\partial (Z^{1/2})^{kj}}{\partial \mu}, \tag{4.12}$$

where μ is the renormalization scale of the theory. Exact results for the superpotential in supersymmetric field theories are important for understanding the dynamics of these theories, and also for investigating dynamical breaking of supersymmetry.

We also note that the instanton induced superpotential which arises in $\mathcal{N}=1$ supersymmetric $SU(N)$ QCD with N_f massless fundamental matter multiplets has been determined exactly [142]. Non-perturbative effects, namely instantons, generate a superpotential in $\mathcal{N}=1$ supersymmetric QCD coupled to massless matter fields, as demonstrated in [130]. The form of this superpotential is fixed up to a numerical coefficient by the requirements of supersymmetry, gauge invariance and global symmetries. The superpotential breaks supersymmetry, and so instantons break supersymmetry in this case [130]. The numerical coefficient of the superpotential, which is non-zero, has been calculated by explicitly determining the superpotential directly from instanton calculations [142], and constitutes an exact non-perturbative result.

The derivations of these exact results are lengthy and would lead us to digress from the theories of primary interest in this thesis, namely $\mathcal{N}=2$ supersymmetric $SU(N)$ gauge theories, and so we do not describe further exact results in $\mathcal{N}=1$ supersymmetric gauge theories here.

Results in supersymmetric gauge theories can sometimes only be found when the theory is in a particular phase. This will be the case in Subsection 4.5.2, where results relating to the vacuum structure of $\mathcal{N} = 1$ supersymmetric quantum Yang–Mills gauge theory in certain phases, made possible by a form of duality, known as Seiberg duality, shall be briefly described.

In this chapter we have not endeavoured to describe the many exact results in $\mathcal{N} = 1$ supersymmetric Abelian gauge theories, such as $\mathcal{N} = 1$ supersymmetric quantum electrodynamics, abbreviated to $\mathcal{N} = 1$ SQED. It is notable that in $\mathcal{N} = 1$ SQED the low energy effective action of the theory is exactly calculable in some cases [143], which anticipates the results for $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory described in Chapter 5. Also in Chapter 5 we will describe the prepotential of $\mathcal{N} = 2$ supersymmetric QCD (in the context of Seiberg–Witten theory), which is significant because it is perturbatively exact but receives non-perturbative corrections, which can be exactly calculated.

4.3 Exact Results in $\mathcal{N} = 4$ Supersymmetric Gauge Theories

For $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory, the exact results often assume the form of a null result. As in any supersymmetric gauge theory, the vacuum energy is precisely zero, as expressed in Eq. (4.1).

Using the NSVZ beta function [136, 139], one can deduce the exact beta function of quantum $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory. The formula for $\beta(g)_{\mathrm{NSZV}}$ given in Eq. (4.5), can be used to derive the beta function of $\mathcal{N} = 4$ supersymmetric Yang–Mills theory by the addition of appropriate matter fields. In practice, this can be done by imposing condition on the $\mathcal{N} = 2$ supersymmetric beta function. For $\mathcal{N} = 4$ theories, one has $\sum_i T(R_i) = 2T_G$, with the following result for the $\mathcal{N} = 4$ beta function:

$$\beta_{\mathcal{N}=4}(g) \;=\; 0. \tag{4.13}$$

This result for $\beta_{\mathcal{N}=4}(g)$ means that the quantum $\mathcal{N} = 4$ theory is scale invariant and ultra-violet finite. The coupling constant g is fixed to be a constant at all energy scales and the physical content of the theory does not change with variations in g. The Wilsonian beta function for this theory also vanishes and receives no quantum perturbative or non-perturbative corrections. This makes the $\mathcal{N} = 4$ an intriguing model of elementary particles due to its finiteness. Theories with $\mathcal{N} = 4$ supersymmetry also exhibit a wealth of correspondences with various string theories; the first such and most significant correspondence of this kind was the 'AdS/CFT' correspondence, a detailed review of which can be found in [277].

Further exact results exist for $\mathcal{N} = 4$ supersymmetric gauge theories, but some of these rely upon the concept of electric-magnetic duality, a property of potentially great physical importance which the theory is conjectured to exhibit. In Section 4.4 below we describe this property and some of the evidence which supports its existence in $\mathcal{N} = 4$ theories.

4.4 Duality in $\mathcal{N} = 4$ Supersymmetric Gauge Theories

In this section we review Montonen–Olive electric-magnetic duality in $\mathcal{N} = 4$ supersymmetric gauge theories. This is a conjectural duality relating electric and magnetic states in gauge theories. The $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory is a candidate theory in which the Montonen–Olive electric-magnetic duality may be realized exactly. It is also the simplest theory which may support the conjecture. In Chapter 5 we shall describe $\mathcal{N} = 2$ supersymmetric gauge theories, in which an effective form of Montonen–Olive electric-magnetic duality may be present.

Duality in field theory concerns the relation between two different types of behaviour in a theory or the relation between two different theories. An example of the latter is the equivalence, at the quantum level, of the sine-Gordon and Thirring models in two spacetime dimensions [191]. In this section we will describe the former kind of duality, in which different regimes of behaviour within the same theory are related, for $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory. From such examples, the following general properties regarding duality can be given:

1. Duality relates the strong coupling regime and the weak coupling regime.

2. Duality interchanges the fundamental (point-like) particle states with solitonic (extended object) states, by which the same theory may be described by different forms of particles. The complete particle spectrum then contains the fields in the classical theory and the soliton states.

3. Duality interchanges Noether currents with topological currents.

In the following subsections we make use of the reviews on duality in field theory [159, 198, 199] and also the reviews [153, 191, 190]. We note the original papers in [155, 156, 157] regarding S-duality in $\mathcal{N} = 4$ supersymmetric gauge theories. We begin first by describing the presence of magnetic monopoles in gauge theories in Subsection 4.4.1 below.

4.4.1 Magnetic Monopoles in Gauge Theories

In this subsection we describe magnetic monopoles in gauge theories. We describe the Dirac monopole [148], the first known magnetic monopole in field theory. The Dirac monopole occurs in quantum electrodynamics if the existence of a magnetic counterpart to the electric current is postulated. Such a postulate requires electric-magnetic duality in the theory, and the Dirac quantization condition relates the electric and magnetic coupling constants in the theory.
We then describe magnetic monopoles in non-Abelian gauge theories, and the first such monopole discovered in these theories, the 't Hoof-Polyakov monopole [149]. The Bogomol'nyi bound [150], a fundamental property of magnetic monopoles and dyons in non-Abelian gauge theories, is then described. Associated with this bound are a new form of particle state, named BPS states [150], which are important states in quantum field theories. We follow the treatment of these topics given in the reviews [153, 191, 190].

Magnetic Monopoles in Abelian Gauge Theories: The Dirac Monopole

The Maxwell equations for classical electromagnetism in four spacetime dimensions can be written as

$$\partial_n v^{mn} = -j^m, \quad \partial_n {}^*v^{mn} = 0, \tag{4.14}$$

where j^m is the four-vector electric current, v_{mn} is the Abelian electromagnetic gauge field strength, defined as $v_{mn} = \partial_m v_n - \partial_n v_m$ and $^*v_{mn}$ is dual of v_{mn} as defined Eq. (2.5) in Section 2.2 of Chapter 2. In vacuo, for which $j^m = 0$, there exists the following duality rotations under which Maxwell's equations for the vector components of v_{mn}, namely the electric and magnetic field strengths, respectively, $\mathbf{E}$ and $\mathbf{B}$ are invariant:

$$(\mathbf{E} + i\mathbf{B}) \ \rightarrow \ (\mathbf{E}' + i\mathbf{B}') \ = \ (\mathbf{E} + i\mathbf{B}) \rightarrow e^{-i\nu}(\mathbf{E} + i\mathbf{B}), \tag{4.15}$$

where ν is an arbitrary angle. When $\nu = \pi/2$, this duality becomes a discrete symmetry, which we label as D, given by:

$$D : \mathbf{E} \rightarrow \mathbf{E}' \ = \ \mathbf{B}, \ \ \mathbf{B} \rightarrow \mathbf{B}' \ = \ -\mathbf{E}. \tag{4.16}$$

The square of the mapping D gives charge conjugation, which we label as C:

$$C \ = \ D^2 : (\mathbf{E}, \mathbf{B}) \rightarrow (\mathbf{E}', \mathbf{B}') \ = \ (-\mathbf{E}, -\mathbf{B}). \tag{4.17}$$

The mapping D is equivalent to a mapping relating the electromagnetic gauge field strength and its dual:

$$D : v_{mn} \rightarrow v'_{mn} \ = \ ^*v_{mn}, \ \ ^*v_{mn} \rightarrow {}^*v'_{mn} \ = \ -v_{mn}. \tag{4.18}$$

This equivalence can only hold in four spacetime dimensions since only in this dimension of spacetime do the electric and magnetic component fields transform as vectors. The duality symmetry of Eq. (4.15) is broken if there is a non-zero current j^m, that is, when matter is present. However, it can be restored by introducing a non-zero purely magnetic current k^m, such that

$$\partial_n {}^*v_{mn} \ = \ k^m. \tag{4.19}$$

When this is done, the discrete symmetry D in Eq. (4.18) acts on the currents as:

$$D : j^m \rightarrow j'^m \ = \ k^m, \ \ k^m \rightarrow k'^m \ = \ -j^m. \tag{4.20}$$

So far this electric-magnetic symmetry is a classical phenomenon. The appearance of a purely magnetic counterpart to the electric current to preserve the discrete electric-magnetic symmetry (electric-magnetic duality) above prompts the introduction of corresponding quanta for the magnetic current, which possess only magnetic charge. Particles which possess free magnetic charge and no other charge are known as magnetic monopoles. In quantum electrodynamics, electric-magnetic duality is present if the electromagnetic vector potential v_m, which now also describes the magnetic monopole, is singular inside the monopole. A consistent solution for the vector potential v_m can be found on the sphere S^2. This solution consists of two vector potentials, one for each hemisphere of S^2, related by a gauge transformation, and both potentials reproduce the same global magnetic field strength. This ensures that the field strength v_{mn} remains continuous and unambiguous.

To be consistent with quantum mechanics, one requires that the phase of the wavefunction describing the magnetic monopole is continous. This imposes the following condition on the unit of magnetic charge g and the unit of electric charge e:

$$eg = 2\pi n\hbar c, \quad n \in \mathbb{Z}, \tag{4.21}$$

which is known as the Dirac quantization condition, and was first derived by Dirac [148]. Thus by assuming the existence of free magnetic charges, the quantization of electric charge is explained and the experimental fact that absolute values of the electron and proton charge are equal is also explained. The quantization condition asserts an inverse relation between the bare electric and bare magnetic running coupling constants of the theory, namely:

$$g = \frac{2\pi n\hbar c}{e}. \tag{4.22}$$

Thus the strong coupling regime in the theory described by electrons is related to the weak coupling theory described by magnetic monopoles. Through this relation, there appears to be two equivalent descriptions of the theory: one in terms of electric quanta, and one in terms of magnetic quanta, which can be transferred between by means (and the consequences of) of the Dirac quantization condition Eq. (4.21).

The Dirac quantization condition in Eq. (4.21) can be generalized to the case of dyons, which are particles possessing both electric and magnetic charges (and no other charge). The result is the Dirac–Zwanziger–Schwinger quantization condition [191], which relates a dyon with electric and magnetic charges (q_1, g_1), respectively, to a dyon with charges (q_2, g_2):

$$q_1 g_2 - q_2 g_1 = 2\pi n\hbar c, \quad n \in \mathbb{Z}, \tag{4.23}$$

restricts the allowed quantized charges of dyons to a two-dimensional lattice. When a magnetic monopole also carries electric charge, and is thus a dyon, the expression of the electric-magnetic duality of Maxwell's equations becomes:

$$e + ig \to e' + ig' \; = \; e^{-i\nu}(e + ig), \tag{4.24}$$

which the Dirac quantization Eq. (4.21) condition does not obey. The Dirac monopole and dyons do not exist in the particle spectrum of quantum electrodynamics. A local theory which describes both electrons and monopoles does not exist.

If electric-magnetic duality is to exist in a Abelian gauge theory, it requires the existence of magnetic monopoles and invariance for all electric and magnetic charges under charge conjugation (the mapping C in Eq. (4.17)). Furthermore, any gauge theory containing a $U(1)$ compact subgroup (which will generate electric charge) of the gauge group will give rise to the possibility of magnetic monopoles in the theory.

Magnetic Monopoles in non-Abelian Gauge Theories: The 't Hooft–Polyakov Monopole

The simplest example of electric-magnetic duality in a non-Abelian (i.e. Yang–Mills) gauge theory is that conjectured for $SO(3)$ or $SU(2)$ Yang–Mills-Higgs gauge theory. The case of $SO(3)$ Yang–Mills-Higgs gauge theory is known as Georgi–Glashow theory [154], and consists of $SO(3)$ Yang–Mills theory coupled to a Higgs triplet field Φ^a. This theory has the following Lagrangian:

$$\mathcal{L}_{\mathrm{GG}} \; = \; -\frac{1}{4}v^a_{mn}v^{amn} + \frac{1}{2}D^m\Phi^a D_m\Phi^a - V(\Phi), \tag{4.25}$$

where v^a_{mn} is the non-Abelian gauge field strength, D^m is the covariant derivative, and $V(\Phi)$ is the Higgs potential given by:

$$V(\Phi) \; = \; \frac{\lambda}{4}(\Phi^a\Phi^a - v^2)^2, \tag{4.26}$$

in which v is the constant vacuum expectation value of the Higgs field Φ^a, and λ is a parameter of the theory.

Static finite-energy solutions to the equations of motion for the Lagrangian $\mathcal{L}_{\mathrm{GG}}$ exist and the theory admits gauge fields with non-zero magnetic charge, that is, magnetic monopoles. These are referred to as 't Hooft–Polyakov monopoles [149], after the authors who jointly discovered them.

In the Georgi–Glashow model, it is the Higgs field which is responsible for the non-zero magnetic charge of the gauge field. This can be seen in the general solution to the equations of motion for the static finite-energy action:

$$\begin{aligned} v^{amn} &= \frac{1}{v}\Phi^a v^{mn}, \\ v_{mn} &= -\frac{1}{ev^3}\varepsilon^{abc}\Phi^a\partial_m\Phi^b\partial_n\Phi^c + \partial_m v_n - \partial_n v_m. \end{aligned} \tag{4.27}$$

The magnetic charge g in this model obeys the Dirac quantization condition except for a factor of two: $eg = 4\pi n\hbar c$. The factor of two arises from the scaling of electric charge in fundamental representations of $SU(2)$.

The appearance of magnetic monopoles in the Georgi–Glashow model [154] is an instance of a more general phenomenon, elucidated by Goddard et. al [151]. Given a gauge theory, be it Abelian or non-Abelian, with a gauge group G which is broken to a subgroup H by the presence of non-zero vacuum expectation values of a Higgs field, the theory will include magnetic monopoles as solutions to the equations of motion [151]. The topology of the Higgs vacuum can be classified by the homotopy group $\pi_2(G/H)$. General Dirac monopole configurations can be constructed from Abelian gauge fields with gauge group H in a suitable way, and the topology of these configurations can be classified via the homotopy group $\pi_1(H)$. Via the isomorphism:

$$\pi_2(G/H) \simeq \pi_1(H), \tag{4.28}$$

it is evident that the Higgs mechanism, by breaking non-Abelian gauge invariance to a smaller gauge invariance, produces magnetic monopoles in the process of doing so. Thus any non-Abelian gauge theories which includes Higgs fields that break non-Abelian gauge invariance shall possess magnetic monopoles in its particle spectrum [149].

The Bogomol'nyi Bound and BPS States

The Georgi–Glashow model [154], defined by the Lagrangian in Eq. (4.25), admits static magnetic monopole solutions with finite mass. When the electric charge of the gauge field in the Georgi–Glashow model is set to zero, $\mathbf{E} = 0$, an expression for the mass of the purely magnetic gauge field, which is the magnetic monopole field, can be obtained. The monopole mass M_{M} is given by:

$$\begin{aligned} M_{\mathrm{M}} &= \int d^3r \left[\frac{1}{2}(\mathbf{B}^a \cdot \mathbf{B}^a + \mathbf{D}\Phi^a \cdot \mathbf{D}\Phi^a) + V(\Phi)\right], \\ &\geq \int d^3r \frac{1}{2}(\mathbf{B}^a \cdot \mathbf{B}^a + \mathbf{D}\Phi^a \cdot \mathbf{D}\Phi^a), \qquad (4.29) \\ &= \frac{1}{2}\int d^3r (\mathbf{B}^a - \mathbf{D}\Phi^a)(\mathbf{B}^a - \mathbf{D}\Phi^a) + vg. \qquad (4.30) \end{aligned}$$

The form of the gauge field strength in Eq. (4.27) and the Bianchi identity $\mathbf{D} \cdot \mathbf{B}^a = 0$ together with Eqs. (4.30,4.30) imply the following bound on the monopole mass; this is the Bogomol'nyi bound [150]:

$$M_{\mathrm{M}} \geq |vg|. \qquad (4.31)$$

The bound in Eq. (4.31) is saturated if and only if the Higgs potential vanishes, $V(\Phi) = 0$, and if the Bogomol'nyi equation holds, which solves the Bianchi identity $\mathbf{D} \cdot \mathbf{B}^a = 0$:

$$\mathbf{B}^a = \mathbf{D}\Phi^a. \qquad (4.32)$$

The Bogomol'nyi equation is a first order equation which implies the second order equations of motion. The limit on monopole mass provided by the Bogomol'nyi bound can be re-expressed in terms of the equations of motion. The resulting equations are known as the Bogomol'nyi-Prasad-Sommerfield (BPS) equations [150], which assume the form:

$$\mathbf{E}^a = 0, \;\; D_0\Phi^a = 0, \;\; \mathbf{B}^a = \pm\mathbf{D}\Phi^a. \qquad (4.33)$$

The BPS limit is the limit $\lambda \to 0$, which gives a vanishing Higgs potential ($V(\Phi) \to 0$) in Eq. (4.26).

In the quantum Georgi-Glashow theory, the quantum version of the classical BPS solution is a new type of particle state, termed a BPS state. The BPS state is not present in the perturbative particle spectrum of the Georgi-Glashow model and is therefore a non-perturbative state. The mass of the BPS state is proportional to the magnetic coupling constant g, and via electric-magnetic duality inversely proportional to the gauge coupling constant e. Hence the BPS state cannot be observed in the weak coupling limit $e \to 0$, since the BPS mass tends to infinity in this limit.

The Bogomol'nyi equation Eq. (4.32) in $\mathbb{R}^3$ (three spacetime dimensions) can be rewritten as the self-dual Yang-Mills equations on $\mathbb{R}^4$ if the Higgs field is identified with the fourth component of the vector potential, $v_4 = \Phi$.

The electric-magnetic duality transformation expressed in Eq. (4.24) implies a generalization of the Bogomol'nyi bound Eq. (4.31) to include dyons. The general mass bound on dyons of charge (q, g) is given by the BPS mass bound:

$$M_{\mathrm{BPS}} \geq v\sqrt{q^2 + g^2}, \tag{4.34}$$

which is invariant under the electric-magnetic duality transformation Eq. (4.24). The BPS mass bound is a universal formula in the theory: it applies equally to fundamental quanta and to solitons such as monopoles and dyons. Semi-classical quantization of solutions of the BPS equations Eq.(4.33) leads to a charge quantization analogous to the Dirac quantization condition Eq. (4.21):

$$q = en_e, \quad n_e \in \mathbb{Z}. \tag{4.35}$$

In Chapter 5 we will describe the rôle of magnetic monopoles and dyons, and exact results concerning them, in $\mathcal{N} = 2$ supersymmetric Yang-Mills gauge theory. We shall also describe how the BPS mass bound arises naturally in $\mathcal{N} = 2$ supersymmetric Yang-Mills gauge theory.

4.4.2 S-Duality in $\mathcal{N} = 4$ Supersymmetric Yang-Mills Gauge Theory

In this subsection we describe a precise conjecture concerning electric-magnetic duality in quantum field theory made by Montonen and Olive [152]. This conjecture explores the consequences of magnetic degrees of freedom in a quantum field theory. This conjecture is useful in understanding the strongly coupled regime of quantum field theories, and thus, the behaviours of quantum field theories in general. We also describe the Witten effect, which is the consequence of the inclusion of ϑ terms for magnetic charges in non-Abelian gauge theories. By incorporating a ϑ term in the Lagrangians describing these non-Abelian gauge theories, the conjectured electric-magnetic duality of Montonen and Olive can be extended to a duality known as S-duality. An exact S-duality may be present in $\mathcal{N} = 4$ supersymmetric Yang-Mills gauge theory, and we briefly describe the evidence for this in the last part of this subsection.

Montonen–Olive Electric-Magnetic Duality

We now describe the conjecture made by Montonen and Olive [152] on electric-magnetic duality in quantum field theory. Given a gauge field theory with gauge group G broken to a subgroup H by the presence of a non-vanishing vacuum expectation value of the Higgs field, Goddard et. al [151] attempted to classify the monopole configurations invariant under the subgroup H. Via group theory arguments, they derived a general charge quantization condition for the above scenario of the gauge symmetry group G breaking to a subgroup H. The unbroken subgroup H must be compact and connected, which are also the criteria applied to gauge groups. If H is the electric group (for example, with $G = SO(3)$ in the Georgi–Glashow model, $H = U(1)$), and H^v is the magnetic group of the theory, under which the monopole is invariant, then these groups are related by their group weight lattices. The dual relation between them is:

$$(H^v)^v \;=\; H. \tag{4.36}$$

For $H^v = SU(N)$, the dual relation Eq. (4.36) gives $(SU(N))^v = SU(N)/\mathbb{Z}_N$.
The Montonen–Olive [152] conjecture is based on this result [151]. The Montonen–Olive conjecture states that [152]:

1. Any gauge theory is characterised by the product $H \times H^v$.

2. There exist two equivalent descriptions of the theory described by the same Lagrangian: (i) one in terms of H-gauge fields with fundamental charged particles in the perturbative spectrum, and (ii) one in terms of H^v-gauge fields with solitons, such as monopoles, in the perturbative spectrum.

A consequence of this is that Noether currents, arising from electric charges, are interchanged with topological currents under the duality relating H and H^v. Then the coupling constant q of the (electric) theory described by H is replaced with the coupling constant g of the (magnetic) theory described by H^v. Via the Dirac quantization condition, one has $g \propto 1/q$, and thus the strongly coupled regime of one theory is related to the weakly coupled regime in the other. By this fact the Montonen–Olive conjecture resists straightforward proof or disproof. The Montonen–Olive electric-magnetic duality conjecture is non-perturbative in nature. Thus the conjecture cannot be proven using perturbation theory alone. Theories in which there are restrictions on the perturbative and non-perturbative behaviour of the theory, so that they can be controlled, provide the most suitable testing ground for the conjecture. Supersymmetric gauge theories are an example of such theories and in particular $\mathcal{N} = 4$ supersymmetric Yang-Mills gauge theory is a candidate theory in which the Montonen–Olive conjecture may be exactly realized.

The Witten Effect

The BPS states present in the quantum Georgi–Glashow model are not indicative of the quantum Georgi–Glashow model being fully invariant under electric-magnetic duality. This is because in the Georgi–Glashow model, all particle state masses receive quantum corrections. Also, there exist massive gauge bosons with spin in the Georgi–Glashow model, and this implies that the monopole states should also have spin; the origin of this in the Georgi–Glashow model is not clear.

Restrictions must be placed on the Georgi–Glashow model and similar theories if they are to describe dyons and also be invariant under electric-magnetic duality. These additional constraints are provided by supersymmetry, and the inclusion of a ϑ-term, which were described in Chapter 2 and Chapter 3, respectively. When a ϑ term is added to the Lagrangian of a Yang-Mills gauge theory, the allowed values of the electric charge in the monopole sector become shifted, and the monopole acquires an electric charge, becoming a dyon. In the Georgi–Glashow model, this electric shift of the magnetic charge can be expressed as:

$$q \; = \; en_e - \frac{e\vartheta}{2\pi}n_m, \quad n_e, n_m \in \mathbb{Z}, \tag{4.37}$$

where the Dirac quantization condition Eq.(4.21) has been used. The phenomenon of magnetic monopoles acquiring one unit of electric charge through a shift of the ϑ angle by $\vartheta \to \vartheta + 2\pi$ is known as the Witten effect [168]. Thus for any Yang-Mills gauge theory Lagrangian with a ϑ term, the magnetic charge of a dyon will always contribute to its electric charge, provided ϑ is non-zero.

S-Duality

The Montonen–Olive electric-magnetic duality described above can be extended to those gauge theories which include a ϑ term in their Lagrangian. The result is a generalized Montonen–Olive duality known as S-duality. In supersymmetric gauge theories, the inclusion of a ϑ term in the Lagrangian will give the most general supersymmetric Yang-Mills term in the theory and the coupling constants g and ϑ can be combined into the complexified coupling constant τ, defined as in Eq. (2.17) of Section 2.2 of Chapter 2:

$$\tau = \frac{\vartheta}{2\pi} + \frac{4\pi i}{g^2}. \tag{4.38}$$

The physical content of the theory will be invariant under the duality transformation:

$$T : \tau \to \tau' = \tau + 1, \tag{4.39}$$

which is the effect of shifting ϑ by 2π on the complex coupling τ. The mapping T in Eq. (4.39) shifts the electric charge q by one unit if there is a magnetic charge of one unit ($n_m = 1$) already present. Hence this is the Witten effect expressed in terms of the complex coupling τ.

Furthermore, the Montonen–Olive electric-magnetic duality can be expressed as the following duality transformation on τ:

$$S : \tau \to \tau' = -\frac{1}{\tau}. \tag{4.40}$$

Together, the duality transformation TS acting on τ can be written as:

$$TS : \tau \to \tau' = \frac{a\tau + b}{c\tau + d}, \quad \{a, b, c, d\} \in \mathbb{Z}, \tag{4.41}$$

where $ad - bc = 1$. The full duality transformation TS is known as S-duality and generates the special linear group of 2×2 matrices, $SL(2, \mathbb{Z})$. The S-duality group $SL(2, \mathbb{Z})$, is also known as the modular group or infinite duality group. This acts on the quantum numbers for electric and magnetic charge, n_e and n_m, respectively, as follows:

$$TS: \begin{pmatrix} n_e \\ n_m \end{pmatrix} \rightarrow \begin{pmatrix} n'_e \\ n'_m \end{pmatrix} = \begin{pmatrix} a & -b \\ c & -d \end{pmatrix} \begin{pmatrix} n_e \\ n_m \end{pmatrix}. \tag{4.42}$$

S-duality can be realized using the representation of n_e and n_m states as points in the complex plane. In terms of the new variables a and a_D, defined by:

$$a \equiv ve, \quad a_D = \tau a, \tag{4.43}$$

so that:

$$q + ig = e(n_e + n_m \tau), \tag{4.44}$$

the mappings T and S, defined in Eq. (4.39), and Eq. (4.40), respectively, can be expressed as:

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} : \begin{pmatrix} a_D \\ a \end{pmatrix} \rightarrow \begin{pmatrix} a'_D \\ a' \end{pmatrix} = \begin{pmatrix} a + a_D \\ a \end{pmatrix}, \tag{4.45}$$

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} : \begin{pmatrix} a_D \\ a \end{pmatrix} \rightarrow \begin{pmatrix} a'_D \\ a' \end{pmatrix} = \begin{pmatrix} -a \\ a_D \end{pmatrix}, \tag{4.46}$$

and the BPS mass bound becomes

$$M_{\mathrm{M}} \geq |a n_e + a_D n_m|. \tag{4.47}$$

The BPS mass bound for a state on the two-dimensional lattice on which the electric and magnetic charges reside, given by Eq. (4.44), is then proportional to the distance from the origin of the state on the electric-magnetic lattice.

S-Duality in $\mathcal{N} = 4$ supersymmetric Yang-Mills theory

We now describe the appearance of S-duality in $\mathcal{N} = 4$ supersymmetric Yang-Mills gauge theory. The connection between Montonen–Olive electric-magnetic duality and $\mathcal{N} = 4$ supersymmetric Yang-Mills gauge theory is apparent from the fact that the equations of motion of the theory admit BPS monopole solutions. Furthermore, the equations of motion are such that any BPS monopole solution can be embedded in them. The $\mathcal{N} = 4$

BPS monopoles belong to an $\mathcal{N} = 4$ (short) multiplet in the theory and can be obtained from the theory perturbatively. Remarkably, the multiplet of $\mathcal{N} = 4$ BPS monopoles is isomorphic to the $\mathcal{N} = 4$ (short) multiplet of vector bosons in the theory. Thus a multiplet of solitonic states is isomorphic to a multiplet of massive fundamental particle states in $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory [159].

The particle states contained in these multiplets all saturate a supersymmetry mass bound which coincides with the BPS mass bound for dyons (itself the generalized Bogomol'nyi mass bound) given by Eq. (4.34).

A further indication that the $\mathcal{N} = 4$ theory is S-dual is that it has modular invariant partition functions; that is, it has partition functions which are modular forms, invariant under the modular group $SL(2, \mathbb{Z})$. This property is required by S-duality, and is satisfied by $\mathcal{N} = 4$ theories [158].

String theory has provided the context for more recent evidence of S-duality in $\mathcal{N} = 4$ theories. Compactified heterotic superstring theory has a low energy effective action which is equivalent to $\mathcal{N} = 4$ supersymmetric Yang-Mills gauge theory. Using this equivalence, it was shown by Sen [156] that S-duality implies bound states of dyons and BPS monopoles in the theory. An existence proof for these bound states in special cases has been given in [156], which further lends support to $\mathcal{N} = 4$ supersymmetric Yang-Mills gauge theory being S-dual (see also the work in [157]).

These properties are perhaps the strongest indication that the $\mathcal{N} = 4$ theory is exactly Montonen–Olive dual. Since the multiplets containing the states dual to each other are isomorphic, the $\mathcal{N} = 4$ theory is, conjecturally, exactly Montonen–Olive self-dual. Without the restrictions on the perturbative and non-perturbative behaviour of the quantum $\mathcal{N} = 4$ theory provided by supersymmetry, which include results such as non-renormalization theorems, deriving these properties may not have been possible.

Observing that the $\mathcal{N} = 4$ theory behaves as an exactly Montonen–Olive self-dual theory does not constitute proof that the theory is self-dual. This is one of the drawbacks of the Montonen–Olive conjecture. Testing and proving the Montonen–Olive conjecture requires knowledge and control of the strongly coupled regime in non-Abelian gauge theories, which is difficult to obtain. This drawback of the conjecture is common to supersymmetric and non-supersymmetric gauge theories.

Supersymmetry is however able to resolve the other two main drawbacks of the Montenon–Olive conjecture. In gauge theories other than supersymmetric Yang–Mills gauge theories, the invariance of the particle spectrum under duality transformations may be broken due to radiative corrections to the Bogomol'nyi bound through renormalization. In the case of $\mathcal{N} = 4$ supersymmetric Yang–Mills theory, the classical Bogomol'nyi bound receives no quantum corrections, and so this cannot occur. Furthermore, in general, if BPS monopoles are to be interpreted as gauge particles, one would expect them to be spin one particles. However, due to their rotational symmetry they appear to be spin zero particles. We will return to the concept of S-duality in the context of $\mathcal{N} = 2$ supersymmetric gauge theories in Chapter 5.

4.5 Duality in $\mathcal{N} = 1$ Supersymmetric Gauge Theories

In Subsection 4.4.2 of Section 4.4 the concept of Montonen–Olive electric-magnetic duality was introduced and evidence for this conjecture in $\mathcal{N} = 4$ supersymmetric Yang-Mills gauge theory was described. In this section we shall briefly outline the work which has uncovered duality in $\mathcal{N} = 1$ supersymmetric gauge theories. This duality in known as Seiberg duality.

The basic statement of Seiberg duality in $\mathcal{N} = 1$ supersymmetric gauge theories is that the low energy effective (or long distance) physics described by these theories for different gauge groups is the same. That is, the choice of gauge group is independent to some extent of the low energy effective dynamics which the theory describes. This is essentially a differently phrased version of the Montonen–Olive conjecture for the case of effective $\mathcal{N} = 1$ gauge theories.

To prepare for our description of Seiberg duality in $\mathcal{N} = 1$ theories, we first briefly describe the exact results pertaining to the quantum and classical moduli spaces of $\mathcal{N} = 1$ theories with gauge group $SU(N)$, in Subsection 4.5.1. Knowledge of the classical and quantum moduli spaces of these theories is useful for understanding the phases, or regimes, of these theories. We briefly describe the phases of non-Abelian (i.e. Yang–Mills) gauge theories and the phases of $\mathcal{N} = 1$ Yang–Mills theories in Subsection 4.5.2. Then in Subsection 4.5.3 we briefly describe Seiberg duality in $\mathcal{N} = 1$ theories.

In analogy to the reviews of Section 4.4 and Section 4.5, we shall describe exact results and their consequences in $\mathcal{N} = 2$ supersymmetric gauge theories in Chapter 5. The results and concepts in $\mathcal{N} = 2$ theories will overlap with the descriptions of $\mathcal{N} = 1$ and $\mathcal{N} = 4$ theories in this chapter. In this section we make use of the reviews [138, 254, 255, 256, 257], and refer to the original papers in [145, 146, 147, 251, 252, 253].

4.5.1 The Moduli Space of $\mathcal{N} = 1$ Supersymmetric Gauge Theories

Following the derivation of exact results for the superpotential in $\mathcal{N} = 1$ supersymmetric gauge theories, the techniques used were applied to more general considerations regarding the space of vacua of these theories. Classically, the $\mathcal{N} = 1$ supersymmetric theories have vacuum valleys, also known as flat directions. These are points or lines in the field space where the potential of the theory is zero (hence the term 'flat' directions) along which exist physically (i.e. gauge) inequivalent vacuum states (ground states). The set of all flat directions in a classical supersymmetric theory can form a continuous manifold, and this is known as the classical moduli space, or classical moduli space of vacua. The moduli space is singular at the point where the number of massless fields increases and the degeneracy of these states cannot be removed by perturbative quantum corrections. That is, the vacuum degeneracy can persist in the quantum theory. Importantly, though, this vacuum degeneracy is sometimes removed, or 'lifted,' by non-perturbative quantum corrections, which can give rise to a superpotential which connects the vacuum states. In some theories, this does not occur, and the quantum moduli space is as degenerate as the classical moduli space. Singularities of the classical moduli space can be smoothed out in the quantum theory.

In this subsection we describe some of the exact results obtained regarding the moduli spaces of $\mathcal{N} = 1$ supersymmetric $SU(N)$ Yang-Mills gauge theory with N_f matter multiplets, also known as $\mathcal{N}=1$ SQCD, following Seiberg [145]. We denote the fundamental and anti-fundamental chiral matter superfields as Q^i and $\tilde{Q}_i$, respectively, with index $i = 1, \ldots, N$.

$\mathcal{N}=1$ Classical Moduli Space

The classical moduli space $\mathcal{M}_c$ of the $\mathcal{N}=1$ theory can be characterized by a constraint on the chiral matter superfields. Classical flat directions in the theory are given by specific matrix values of the superfields Q^i and $\tilde{Q}_i$. The constraints defining the moduli space take the form of conditions on gauge invariant combinations of the matter superfields. For the $SU(2)$ theory with N_f matter multiplets, the constraints for $N_f \geq 2$ are:

$$V^{ij} = Q^i Q^j, \tag{4.48}$$

$$\epsilon_{i_1,\ldots,i_{2N_f}} V^{i_1 i_2} V^{i_3 i_4} = 0, \tag{4.49}$$

where V^{ij} is a gauge invariant combination of the matter superfields and ϵ is the totally antisymmetric tensor formed from the indices $i_1, \ldots, i_{2N_f}$, where $i, j, k = 1, \ldots, N$ are gauge group indices. When V in Eq. (4.48) assumes a non-zero value, the global gauge symmetry of the theory is completely broken.

For $SU(N)$ theories with $N > 2$, the classical moduli space can be specifed in a similar way, as the constraints:

$$M^i_k = Q^i \tilde{Q}_k, \tag{4.50}$$

$$B_{i_{N+1} i_{N+2},\ldots,i_{N_f}} = \frac{1}{N!}\epsilon_{i_1,\ldots,i_{N_f}} Q^{i_1} Q^{i_2} \cdots Q^{i_{N_f}}, \tag{4.51}$$

$$\tilde{B}^{k_{N+1} k_{N+2},\ldots,k_{N_f}} = \frac{1}{N!}\epsilon^{k_1,\ldots,k_{N_f}} Q_{k_1} Q_{k_2} \cdots Q_{k_{N_f}}. \tag{4.52}$$

These gauge invariant combinations can be considered as particular matter superfields. For differing values of N_f the form of the classical moduli space differs. Furthermore, different points on the moduli space $\mathcal{M}_c$ possess different global and gauge symmetries. A generic point on the moduli space will have an unbroken $SU(N_f - N)$ gauge symmetry, and there exist singular points when $B = \tilde{B} = 0$. From such information one can deduce the structure of the classical moduli space for given values of N and N_f.

$\mathcal{N} = 1$ *Quantum Moduli Space*

Information concerning the quantum moduli space for $\mathcal{N} = 1$ theories is less readily obtained. One expects the classical moduli space $\mathcal{M}_c$ to receive quantum corrections and using the Wilsonian effective action of the theory is one way in which the quantum moduli space $\mathcal{M}_q$ can be explored. Mass terms of the form $m^i_k Q^i \tilde{Q}_k$ can be added to the $\mathcal{N} = 1$ superpotential $\mathcal{W}$ to control the theory along the flat directions of the potential. The resulting theory then has massive matter multiplets provided $m^i_k \neq 0$.

When $\det m^i_k \neq 0$ all of the flat directions of the theory are lifted by quantum corrections and thus are no longer flat. Classically this would imply $M = B = \tilde{B} = 0$. These values are modified by quantum corrections in the full quantum theory.

When $\det m^i_k = 0$, the classical quantities become expectation values and the quantity M in Eq. (4.50) is modified to:

$$M^i_k = \left\langle Q^i \tilde{Q}_k \right\rangle = \Lambda^{3N-N_f/N} (\det m^i_k)^{1/N} \left(\frac{1}{m}\right)^i_k, \tag{4.53}$$

where Λ is the low energy scale and the expectation value for M is exact.

For $N_f < N$, an exact superpotential can be found using the criteria given by Seiberg in Section 4.2. For $N_f \geq N$, the classical vacuum degeneracy of the theory is not lifted and the quantum moduli space consists of these ground states. These results for the massive theory, however, do not extend to the case of zero mass, when $m^i_k = 0$. This is because a non-trivial quantum moduli space, denoted $\mathcal{M}_{q,m\to 0}$ exists in the limit $m^i_k \to 0$. The Wilsonian effective action of the theory will include all of the massless fields in the theory and can be expected correctly give the low energy dynamics of the 'light' fields which are not integrated out of the Lagrangian. It may not give the correct dynamics for massive fields since some of these will have been integrated out of the Lagrangian (as they will be 'heavy' fields which possess mass greater than the energy scale Λ above which fields and modes are integrated out). Thus it is important to consider the $\mathcal{N} = 1$ theory with massless chiral matter multiplets.

In the work of Seiberg, $\mathcal{N} = 1$ SQCD with massless matter multiplets is considered for various values of N_f. For $N_f = N$, the quantum theory is modified such that the quantum space $\mathcal{M}_{q,m\to 0}$ resembles the classical space $\mathcal{M}_c$ except at the classical singularities. The quantum space has a singularity associated with the appearance of massless, or 'light,' fields which may be present on $\mathcal{M}_{q,m\to 0}$. These new massless fields arise when the superfield combinations $V, M, B, \tilde{B}$ become massless. Collectively, these massless superfields are the moduli of the theory, which parameterize the moduli space.

At the points where massless fields emerge, the global symmetry of the theory remains unbroken, and may be enhanced. Since the quantum moduli space $\mathcal{M}_{q,m\to 0}$ is different to the classical moduli space $\mathcal{M}_c$, the low energy effective theory may have soliton states, which are not present on the classical moduli space.

For $N_f = N + 1$, the quantum moduli space is the same as the classical moduli space: $\mathcal{M}_{q,m\to 0} \simeq \mathcal{M}_c$, except at singularities, where other 'light' fields may also be present on $\mathcal{M}_{q,m\to 0}$.

In both the cases of $N_f = N$ and $N_f = N + 1$, the quantum $\mathcal{N} = 1$ superpotential has equations of motion which serve to define the moduli space.

For $N_f \geq N + 2$, again the quantum moduli space is equivalent to the classical moduli space, except at singularities. Massless matter superfields only become possible if the theory is scale invariant, which it is for certain ranges of N_f and N. The exact nature of the quantum moduli space in this case is uncertain due to the interpretation of singular points on $\mathcal{M}_{q,m\to 0}$.

Perturbations on the quantum moduli space can also reveal more about the structure of the massless quantum theory. Such perturbations can be acheived by adding mass terms and additional superfield terms to the $\mathcal{N} = 1$ superpotential. Then vacua and singular points on the moduli space can be identified. There can exist a set of inequivalent discrete vacua in addition to the set of continuous inequivalent ground states in the theory.

In summary, the quantum moduli space for $\mathcal{N} = 1$ SQCD can be elucidated almost completely for the low energy Wilsonian effective theory. The interpretation of singular points on the quantum moduli space, and how classical singular points become modified in the quantum space, are important problems in attempts to comprehend the structure of the moduli space of the quantum theory. In Chapter 5 we will describe a similar analysis for $\mathcal{N} = 2$ supersymmetic Yang-Mills gauge theory and $\mathcal{N} = 2$ SQCD, and the exact results which follow from this.

4.5.2 Phases of $\mathcal{N} = 1$ Supersymmetric Gauge Theories

In general, gauge theories (both Abelian and non-Abelian) exist in a number of dynamical regimes referred to as phases. A phase is essentially a space specified by the parameters in a theory when they assume certain values or behaviour. Variations in the parameters of the theory will induce transitions between these phases. The choice of the vacuum state also dictates in which phase the theory is in.

In this subsection we describe the phases of non-Abelian gauge theories in general and then specialise to the case of the phases of $\mathcal{N} = 1$ supersymmetric gauge theories. In supersymmetric field theories, there can exist many inequivalent vacua for the same set of parameter values. Each of the vacua may be in a different phase, and so each may form the ground state for the theory in a different phase. This indicates the significance of phases in supersymmetric gauge theories.

Phases of non-Abelian Gauge Theories

Yang-Mills or non-Abelian gauge theories exist in four phases. These phases, except for the confining phase, are also present in Abelian gauge theories. These phases are the free or Landau phase, the Higgs phase, the Coulomb phase and the confining phase. The four phases of non-Abelian gauge theories are characterized in general terms by the following properties [138].

Free phase: occurs when the mass parameter vanishes and the scalar fields of the theory have vanishing vacuum expectation values. Any long-range potentials are screened by quantum effects. Massless gauge bosons become 'dressed' by virtual particles. The asymptotic limit of the massless theory results in a free gauge boson and massless matter fields. The matter fields have their electric (i.e. Abelian) charges completely screened by vacuum condensates of scalar fields. No conventially defined S-matrix exists and there are no localized asymptotic states. The theory is ill-defined at short distances. Also known as the Landau zero-charge phase.

Higgs phase: occurs when the mass parameter vanishes and the scalar fields of the theory assume non-vanishing vacuum expectation values. The gauge symmetry is spontaneously broken. Vector (boson) particles acquire mass via the Higgs mechanism. The electric charge of fields is screened by vacuum condensates. The free phase is a special point (often the origin in field space) in the Higgs phase where the gauge symmetry is unbroken.

Coulomb phase: occurs when the mass parameter is non-zero and the scalar fields of the theory have vanishing expectation values. The vacuum state is non-degenerate and thus unique. The potential for interaction between static electric charges tends towards a Coulombic potential.

Confining phase: occurs only in Yang-Mills gauge field theories. This is a more complicated phase. The potential for interaction between static non-Abelian (i.e. colour) charges tends towards a linear function of the charge separation. Thus the long-range force between the colour charges increases with increasing distance and the charges cannot be separated at asymptotically large separations (as this would require infinite energy). Thus the colour charges are confined. Only at asymptotically short distances can they be separated; this is the phenomenon of asymptotic freedom. The Coulombic field of the Coulomb phase is replaced by a flux tube formed by the colour charge field.
The Meissner effect in the Ginzburg-Landau theory of superconductivity exhibits (non-relativistically) analogous behaviour. In the Meissner effect, magnetic charges are repelled from the vacuum by the formation of magnetic flux tubes which connect them. Thereby the magnetic charges form a condensate in the vacuum. In the confining phase of a Yang-Mills gauge theory, colour charges are repelled by the vacuum and are connected by flux tubes. Magnetic monopoles are theorized to condense in these theories and to induce the formation of colour flux tubes and prevent electric charges condensing.
Thus there apparently exists a colour analogue of the Meissner effect in the theory; this is known as the dual Meissner effect. However, Yang-Mills gauge theories are strongly coupled theories and whether a dual Meissner effect takes place in these theories, and particularly in QCD, remains inconclusive. In Chapter 5 we shall describe the dual Meissner effect and monopole condensation in $\mathcal{N} = 2$ supersymmetric gauge theories.

In addition to these phases, there also exist variants of the Higgs phase and the confining phase. These are the unified Higgs/confining phase and the oblique confining phase. The oblique confining phase is of some relevance to supersymmetric Yang-Mills gauge theories.

Oblique confinement occurs when dyons, generically particles possessing both non-zero magnetic and non-zero electric charge, condense. Dyons necessarily exist in a theory which has magnetic monopoles, since a non-zero vacuum angle ϑ universally connects magnetic and electric charge. Unfortunately, some fractional charges (as dyon-quark bound states) appear in the particle spectrum in this phase. This phase remains a possible first step in attempts to decide if a dual Meissner effect occurs in QCD and other Yang-Mills gauge theories.

Phases of $\mathcal{N} = 1$ Supersymmetric Gauge Theories

In $\mathcal{N} = 1$ supersymmetric Yang-Mills theory, the phases present include all of those which occur for non-Abelian gauge theories in general [145, 147]. However, supersymmetry constrains the form of the moduli spaces in each phase. The $\mathcal{N} = 1$ superpotential $\mathcal{W}$ is holomorphic in the moduli (massless matter superfields) and the coupling constants. This holomorphy forbids any first order phase transitions. In $\mathcal{N} = 1$ theories there are two classes of phases:

1. The moduli space of a phase possesses a moduli subspace which is smaller than the moduli space and is of a size greater than or equal to the moduli space of another phase.

2. Different phases of the theory overlap in the moduli space of theory at transition points.

The confining phase and Higgs phase of $\mathcal{N} = 1$ theories are not distinguishable when the matter multiplets are in the fundamental representation. Such theories do not have ground states in the Coulomb phase. Using the techniques pioneered by Seiberg and Witten for $\mathcal{N} = 2$ supersymmetric gauge theories, precise statements about the nature of singularities on the quantum moduli space of some $\mathcal{N} = 1$ theories can be made. A conjectured generic feature of singular points on $\mathcal{N} = 1$ quantum moduli spaces is that these points are due to the appearance of massless magnetic monopoles. When the quantum moduli space is perturbed, these magnetic monopoles condense and the theory enters a confining phase. We will describe the results for the quantum moduli space of $\mathcal{N} = 2$ supersymmetric gauge theories in Chapter 5. These results can also be extended to $\mathcal{N} = 1$ theories and enable one to determine the form of the complex coupling constant τ for low energy Wilsonian effective $\mathcal{N} = 1$ supersymmetric gauge theories.

In the Coulomb phase, $\mathcal{N} = 1$ theories possess massless photons and may possess Seiberg duality. In Subsection 4.5.3 below we describe Seiberg duality in $\mathcal{N} = 1$ SQCD in terms of $\mathcal{N} = 1$ phases and moduli spaces.

4.5.3 Seiberg Duality in $\mathcal{N} = 1$ Supersymmetric Gauge Theories

In Section 4.2 one of the exact results which can be obtained in some quantum $\mathcal{N} = 1$ supersymmetric gauge theories is the exact effective $\mathcal{N} = 1$ superpotential, which we denote $\mathcal{W}_{\text{eff}}$. We described how a small number of constraints was sufficient in many cases to exactly determine the form of $\mathcal{W}_{\text{eff}}$ [141]. In Section 3.3 of Chapter 3 it was explained that the superpotential in $\mathcal{N} = 1$ theories assumes a highly important rôle in determining the dynamics of the theory. Given knowledge of the effective superpotential, the dynamics and behaviour of light particles can be elucidated. Furthermore, information regarding the phase structure and phase transitions of the theory can be then be extracted.

In this subsection we describe the results on Seiberg duality in quantum $\mathcal{N} = 1$ supersymmetric QCD (SQCD): that is, $\mathcal{N} = 1$ supersymmetric $SU(N)$ gauge theories with N_f matter multiplets as described Section 3.3 in Chapter 3. These results have been obtained by Seiberg and co-workers [145, 146, 147, 251, 252] through the process of determining the effective superpotential exactly and following the implications of the form of $\mathcal{W}_{\text{eff}}$ for the theory. In this subsection we follow the reviews in [254, 255, 256, 257] and the original papers [145, 146, 147, 251, 252].

In $\mathcal{N} = 1$ SQCD with gauge group $SU(N)$, the phase of the theory and the phenomenon it exhibits is dependent on the number of massless quarks, N_f, which appear in the N_f chiral matter multiplets in the fundamental representation of the theory. The Lagrangian for $\mathcal{N} = 1$ SQCD was given in Eq.(3.79) of Section 3.3 of Chapter 3. We now describe the ranges for N_f and the phenomenon which arise for these values of N_f.

$N_f \geq 3N$

The theory is in the free phase. To be precise, it is the free phase for the electrically charged particles in the theory, corresponding to the ordinary free (Landau) phase. The particle spectrum at large distance consists of quarks and gluons. For a separation R, the long distance behaviour of the electric potential has the form:

$$V \sim \frac{1}{R \log R}, \tag{4.54}$$

which decreases faster with R than a Coulombic potential. Although in this phase the theory is not a well defined quantum field theory, as described in Subsection 4.5.2, this theory can be a consistent low energy description of another theory.

$\frac{3}{2}N < N_f < 3N$

The theory is in the Coulomb phase and is asymptotically free. The coupling constant does not increase indefinitely, but reaches a fixed finite value at large distances: this is a fixed point of the renormalization group. The theory is now therefore a non-trivial four dimensional conformal quantum field theory. The long distance electric potential behaves as:

$$V \sim \frac{1}{R},$$

which is a Coulombic potential. At the fixed value of the coupling constant the quarks and gluons in the theory appear as massless interacting particles and are not confined. It has been shown by Seiberg that there exists a magnetic description of $\mathcal{N} = 1$ SCQD with gauge group $SU(N_f - N)$ and N_f matter multiplets at this value of the coupling constant. We can express this Seiberg duality schematically as:

$$\mathcal{N} = 1 \text{ SCQD} \; : \; SU(N) \text{ (electric)} \longleftrightarrow SU(N_f - N) \text{ (magnetic)}.$$

The dual theory is also in the Coulomb phase. The particle states in the dual theory will possess magnetic charge and the dual gauge group $SU(N_f - N)$ can be interpreted as the 'magnetic' gauge group. The original gauge group $SU(N)$ is then the 'electric' gauge group. The low energy (or equivalently, long distance) physics of the electric $SU(N)$ theory is identical to the low energy physics of the magnetic $SU(N_f - N)$ theory. This is a remarkable fact given that the theory and its dual are invariant under different gauge groups and describe different numbers of interacting particles.

The fixed point in the original theory appears in the dual theory and is described by variables of both interacting electric and magnetic states. Experimentally there would be no way to determine whether the Coulombic potential in this phase is mediated by the electric or magnetic states. As N_f is decreased, the magnetic gauge group becomes smaller and so the magnetic theory becomes weaker via the Higgs mechanism. In comparison, the electric theory becomes stronger as N_f decreases. Physically, a reduction in N_f can be achieved by assigning masses to quarks and then decoupling the quarks.

At low energies in the theory, the strong coupling regime involving electric particles can be described using the weak coupling regime involving magnetic particles. In this range of N_f, therefore, there is evidence which indicates the existence of Seiberg duality in $\mathcal{N} = 1$ supersymmetric $SU(N)$ Yang-Mills gauge theory with N_f matter multiplets.

$N + 2 \leq N_f \leq \frac{3}{2}N$

In this range of N_f, the theory is in the free phase for magnetically charged particles. This is because the electric theory is very strongly coupled and so the dual description of the theory in terms of magnetic particles becomes very weakly coupled to the extent that it is a free theory. This phase occurs as the magnetic $SU(N_f - N)$ theory is not asymptotically free and is weakly coupled at low energies. The magnetic massless particles are composites of the electric fundamental states and exhibit a gauge invariance (under $SU(N_f - N)$) which is not manifest in the electric description of the theory. Hence, through duality, a new gauge invariance of the theory (in its dual form) is revealed. The form of the potential between magnetic particles is the same as that for the free electric phase ($N_f \geq 3N$) in Eq, (4.54).

$N_f = N + 1$ and $N_f = N$

In this range the process of reducing N_f by decoupling quarks via assigning them masses results in the eventual complete breaking of the magnetic gauge group $SU(N_f - N)$. During this reduction, the Higgs mechanism continues to provide mass for the gauge boson of the theory, until there are no massless gauge particles present. This serves to confine the electric particles completely, and so this phase can be interpreted as the confining phase. Then baryons in the theory, inside which electric states such as quarks and gluons are confined, become magnetic monopoles.

$N_f < N$

This is a range of N_f for which the theory has no ground state. The massless quarks of the theory possess no lowest energy state, making the theory unphysical.

We note that a similar form of Seiberg duality is exhibited by $\mathcal{N} = 1$ supersymmetric $SO(N)$ Yang-Mills gauge theory with N_f matter multiplets in the fundamental representation of the gauge group $SO(N)$. The duality map in this case between electric and magnetic states is:

$$SO(N) \text{ (electric)} \longleftrightarrow SO(N_f - N + 4) \text{ (magnetic)},$$

where the magnetic $SO(N_f - N + 4)$ theory has N_f matter multiplets in the fundamental representation of the dual gauge group $SO(N_f - N + 4)$. This theory also has an oblique confining phase which can be equivalently described by dyons, and so there exists an electric-magnetic-dyonic triality in this theory.

In $\mathcal{N} = 1$ supersymmetric gauge theories, as with other gauge theories which realize electric-magnetic duality, the electric degrees of freedom are related non-locally to the magnetic degrees of freedom of the theory. The magnetic theory can be considered as an effective theory which describes the low energy (i.e. long distance) electric theory.

Furthermore, through Montonen–Olive duality, the magnetic theory provides a weakly coupled description of the strongly coupled electric theory, which is useful in exploring strongly coupled phases such as the confining phase. In this way, hitherto unknown phases of $\mathcal{N} = 1$ supersymmetric gauge theories are found, such as the free magnetic phase and the non-Abelian Coulomb phase in $\mathcal{N} = 1$ SQCD.

In Chapter 5 we shall describe exact results obtained in four dimensional $\mathcal{N} = 2$ supersymmetric gauge theories. In analogy to the results given here for $\mathcal{N} = 1$ theories, exact results for the $\mathcal{N} = 2$ superpotential and beta function shall be described. The exact results for the low energy effective $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang-Mills gauge theory, recently proposed by Seiberg and Witten, known as Seiberg–Witten theory, and its generalizations, will also be described.

Chapter 5

Exact Results in Supersymmetric Gauge Theories II: $\mathcal{N} = 2$ Supersymmetry

5.1 Introduction

In this chapter we describe exact results in $\mathcal{N} = 2$ supersymmetric gauge theories. The most recent of these are the proposed exact solutions for the low energy Wilsonian effective actions of $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory with various gauge groups and matter content. The techniques used in these models to determine their low energy effective actions are based upon the pioneering work of Seiberg and Witten [170, 171]. These authors were able to propose, through an elaborate chain of reasoning, the exact low energy effective action for both $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory [170] and $\mathcal{N} = 2$ $SU(2)$ SQCD [171], in 1994. The proposed exact results for the low energy effective actions of these two particular models are known as Seiberg–Witten theory [170, 171]. The status of these purportedly exact results is essentially that of conjecture, but there exists compelling evidence in favour of their results from many calculations and checks. One such set of checks are those involving instanton methods, which we describe in Chapter 6.

The amount of scientific literature regarding Seiberg–Witten theory is vast, and our description will inevitably be incomplete. Seiberg–Witten theory is a major development in quantum field theory and many new lines of enquiry and results stem from their work, extending beyond supersymmetric field theories to string theory and pure mathematics. The formalism for supersymmetric gauge theories in Chapter 3 will be used, and we also utilize many of the notions which were introduced in Chapter 4 in the context of $\mathcal{N} = 1$ supersymmetric gauge theories. We continue to work in four dimensional Minkowski spacetime, whose metric is given in Appendix A.

We first describe some of the exact results determined for $\mathcal{N} = 2$ supersymmetric gauge theories in Section 5.2. These results were found before the advent of Seiberg–Witten theory. In Section 5.3 we describe Seiberg–Witten theory. We devote one subsection each to the results derived by Seiberg and Witten, for $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, and $\mathcal{N} = 2$ $SU(2)$ SQCD. In the latter subsection, we also briefly note other results in Seiberg–Witten theory. Generalizations of Seiberg–Witten theory to models with different gauge groups, including $SU(N)$, are described in Section 5.4. We also describe some of the generalizations of Seiberg–Witten theory to models with exceptional gauge groups and other matter content, and describe other results relating to Seiberg–Witten theory, in Section 5.4.

5.2 Exact Results in $\mathcal{N} = 2$ Supersymmetric Gauge Theories

In this section we briefly describe some of the exact results and related results obtained in $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory and $\mathcal{N} = 2$ SQCD. This will be analogous to Section 4.2 of Chapter 4, in which exact results in $\mathcal{N} = 1$ supersymmetric gauge theories were described.

As in any supersymmetric field theory, the first exact result which can be readily obtained from $\mathcal{N} = 2$ theories is the vanishing of the vacuum energy $E_{\text{vac}} = 0$. This vacuum energy E_{vac} receives no perturbative or non-perturbative quantum corrections, and is thus exact to all orders in the gauge coupling constant. The reason for this result and an indication of the derivation of this result are given in Section 4.2 of Chaper 3.

In the following paragraphs we shall briefly outline some of the other exact results obtained in $\mathcal{N} = 2$ theories. We make use of the reviews in [190, 191] and refer to the original papers in [160, 161, 162, 163, 164, 165, 166, 169].

Masses and Electric Dipole Moments of Supersymmetric Monopoles

The occurrence of magnetic monopole and dyon solutions to the field equations of supersymmetric Yang–Mills gauge theories coupled to a scalar field has been demonstrated in [160]. The non-zero vanishing expectation of the scalar field in these theories spontaneously breaks the gauge group and permits soliton solutions of the classical field equations to exist. An example of a supersymmetric field theory containing magnetic monopoles and dyons is the $\mathcal{N} = 1$ supersymmetric version of the Georgi–Glashow model introduced in Chapter 4. As has been described in Chapter 4, the presence of supersymmetry ensures the cancellation of certain quantities associated with bosonic and fermionic states in a quantum field theory. As will be described in Chapter 6, such cancellations occur in semi-classical calculations about instanton configurations in supersymmetric gauge theories, and consequently simplify such calculations.

The cancellation of bosonic and fermionic contributions to quantities relating soliton solutions in $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theories coupled to a scalar field has consequences for the masses of the soliton or BPS states. The mass of any classical solution of the field equations of these supersymmetric field theories do not receive quantum perturbative one loop corrections [160]. However, they may still receive quantum non-perturbative corrections. Let the classical mass of a magnetic monopole or dyon solution in an $\mathcal{N} = 2$ Yang–Mills gauge theory coupled to a massive scalar field ϕ be M_{BPS}. Then in the full quantum theory the one loop quantum corrected soliton mass M^q_{BPS} is given by:

$$M^q_{BPS} = M_{BPS}. \tag{5.1}$$

This perturbatively and non-perturbatively exact result is significant in $\mathcal{N} = 4$ and $\mathcal{N} = 2$ supersymmetric field theories as it can be interpreted as evidence supporting the conjectured Montonen–Olive electric-magnetic duality or S-duality in these theories.

A further exact result regarding magnetic monopoles in supersymmetric gauge theories is the calculation of the electric dipole moment for the monopole [161]. The electric dipole moment for magnetic monopoles in a quantum field theory arises from the spin of the fermion zero modes of the monopole. It is analogous to the electric dipole moment for electrically charged particles. The magnitude of the electric dipole moment ν has a magnitude defined by:

$$\nu = -g_M \left(\frac{g}{2M}\right) \mathbf{S}, \tag{5.2}$$

where $\mathbf{S}$ is the spin and g is the magnetic charge of the monopole, which has mass M. The quantity g_M is referred to as the magnetic 'g-factor', which is defined by Eq. (5.2). The spin $\mathbf{S}$ is a quantum effect and the electric dipole moment for the monopole may be calculated semi-classically. By considering the response of the monopole to a weak constant external electric field, the factor g_M for the monopole can be explicitly determined [161]. The exact result for g_M is:

$$g_M = 2, \tag{5.3}$$

which agrees with the 'g-factor' in the electric dipole moment for elementary electrical states, $g_E = 2$. This result is in accordance with the Dirac equation and also for massive spin 1 particles created via the Higgs mechanism in spontaneously broken gauge theories. For $\mathcal{N} = 4$ supersymmetric gauge theories, this result can also be interpreted as evidence supporting Montonen–Olive electric-magnetic duality.

Renormalization of Multiplet Masses

The divergent renormalization constants in general supersymmetric Yang–Mills gauge theories have been evaluated to one loop order in [162]. With particular application to $\mathcal{N} = 2$ supersymmetric gauge theories, a new renormalization theorem for the masses of matter multiplets was uncovered through the work in [162].

For $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory with simple gauge group G coupled to matter multiplets of arbitrary mass and transforming in general representations of G, the mass of the matter multiplets is unrenormalized. Quantitatively, if the renormalized mass of a $\mathcal{N} = 2$ matter multiplet is m_{ren}, and m_{bare} is the bare, unrenormalized mass of the same matter multiplet, then one has:

$$m_{ren} = m_{bare}, \tag{5.4}$$

which holds to all orders in the gauge coupling constant g. Thus the bare matter multiplet mass receives no quantum perturbative or non-perturbative corrections. This exact result is valid due to the presence of $\mathcal{N} = 2$ supersymmetry, and follows from the non-renormalization of the BPS mass (M_{BPS}) related to the $\mathcal{N} = 2$ central charge.

Wilsonian Beta Function

The beta function of the Wilsonian effective $\mathcal{N} = 2$ supersymmetric Yang—Mills gauge theory, with or without matter fields, denoted by $\beta^W_{\mathcal{N}=2}(g)$, can be exactly determined. The result for the Wilsonian beta function in $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills theory is one loop exact and can be obtained via perturbative calculations. The result is:

$$\beta^W_{\mathcal{N}=2}(g) = -\frac{g^3 N(4 - N_f)}{32\pi}. \tag{5.5}$$

This expression for the Wilsonian beta function $\beta^W_{\mathcal{N}=2}(g)$ is exact: it receives no perturbative or non-perturbative quantum corrections.

It has also been shown that under certain conditions the full $\mathcal{N} = 2$ beta function $\beta_{\mathcal{N}=2}(g)$ vanishes in $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theories coupled to $\mathcal{N} = 2$ matter multiplets [163]. The resulting $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory is finite and scale invariant. This is an exact result pertaining to the $\mathcal{N} = 2$ supersymmetric beta function which can be derived from the NSVZ beta function when $\mathcal{N} = 2$. The $\mathcal{N} = 2$ beta function vanishes when there are a specific number of $\mathcal{N} = 2$ matter multiplets to which the $\mathcal{N} = 2$ Yang–Mills gauge theory is coupled, transforming in a particular representation of the gauge group G of the theory.

To state the conditions for which the $\mathcal{N} = 2$ beta function vanishes, we take the gauge group G to be $G = SU(N)$. Then let the masses of the $\mathcal{N} = 2$ matter multiplets be m_i, where $i = 1, \ldots, N$. Furthermore, let the matter multiplets transform in representations of G denoted by R_i. By considering the general group theoretic expression for the $\mathcal{N} = 2$ beta function when there are N_f massive matter multiplets present, a solution for which

$\beta_{\mathcal{N}=2}(g) = 0$ is given by:

$$N_f = 2N, \tag{5.6}$$

$$N \sum_i T(R_i) = C_2(G), \tag{5.7}$$

where C_2 is the second Casimir of the gauge group G, and $T(R_i)$ are group factors of the representations R_i the group G. Hence if there are an even number of matter multiplets in representations R_i of G which obey Eq. (5.7), the $\mathcal{N} = 2$ beta function is exactly zero to all orders in the gauge coupling constant g, and the theory is also finite to all orders in g.

5.3 Seiberg–Witten Theory

In this section we make use of the original papers by Seiberg and Witten [170, 171], the lectures by Witten [172] and some of the many reviews on Seiberg–Witten theory, which include [190, 191, 193, 194, 192, 195, 196, 197, 248], and those reviews which include material on Seiberg–Witten theory, for example [198, 199, 200, 201]. In particular, we follow the reviews [190, 191, 193].

The exact result for the low energy Wilsonian effective action of $\mathcal{N} = 2$ supersymmetric Yang–Mills theory proposed by Seiberg and Witten is the first exact solution for the low energy dynamics of a four dimensional quantum field theory. This is a significant result in the study of four dimensional quantum field theory, the class of which phenomenologically important theories such as quantum electrodynamics (QED) and quantum chromodynamics (QCD) belong. Seiberg and Witten were able to give a complete description of the vacuum structure, that is, the moduli space, of $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory in their work, which enables the reconstruction of the low energy Wilsonian effective action of this theory.

The presence of supersymmetry permits control over the quantum corrections and the calculation of exact results in the theory. The property of holomorphy is also vital for their analysis, as is a form of electric-magnetic duality. Combined with physical intuition and reasoning, these concepts lead to Seiberg–Witten theory.

In Subsection 5.3.1 we describe the work in the first paper of Seiberg and Witten on the low energy Wilsonian effective action of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills theory [170]. The techniques used in this paper were then extended to $\mathcal{N} = 2$ $SU(2)$ SQCD in [171], which we describe in Subsection 5.3.2.

5.3.1 $\mathcal{N} = 2$ Supersymmetric $SU(2)$ Yang–Mills Gauge Theory

The Lagrangian for this theory in terms of $\mathcal{N} = 2$ superfields with gauge group $SU(2)$ is given by Eq. (3.86) of Section 3.4 of Chapter 3. However, we shall be interested in the low energy Wilsonian effective action of this theory. When the scalar field ϕ in the $\mathcal{N} = 2$ field multiplet acquires a non-zero vacuum expectation value $\langle\phi\rangle$, the gauge group (rank r) of the theory is spontaneously broken to an (unbroken) subgroup $U(1)^r$ of the gauge group. The theory is then in the Coulomb phase. At low energies, the effective theory will contain fewer massive states than massless states, and at sufficiently low energies the physical states of the theory will only be massless states. Thus the effective theory will only describe massless fields (and modes). If the vacuum expectation values $\langle\phi\rangle$ are not degenerate, then the only massless fields present will be those invariant under the unbroken $U(1)^r$ subgroups. Since the $U(1)$ gauge group is associated with electric phenomena, these massless fields will possess electric charge.

As described in Section 4.2 of Chapter 4, there exists a procedure to obtain the low energy Wilsonian effective action of a given field theory. This involves integrating out all massive and massless fields and modes above a given dynamically generated scale Λ, which acts as a low energy cut off scale. In $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory, some results have been obtained for the effective theory [164]. A complete determination of the form of the low energy effective theory was not possible until the work of Seiberg and Witten [170]. Their technique is indirect, and uses pioneering but unconventional methods.

Central Charges in $\mathcal{N} = 2$ Supersymmetric Yang–Mills Gauge Theory

The $\mathcal{N} = 2$ supersymmetry algebra with central charges (and no mass terms) present can be simplified by skew-diagonalising the central charge matrices Z^{IJ}. The $\mathcal{N} = 2$

supersymmetry algebra with central charges than takes the form:

$$\{Q^a_\alpha, \bar{Q}_{\dot\beta b}\} = 2\sigma^m_{\alpha\dot\beta} P_m \delta^a_b, \tag{5.8}$$

$$\{Q^a_\alpha, Q^b_\beta\} = 2\sqrt{2}\epsilon_{\alpha\beta}\epsilon^{ab} Z, \tag{5.9}$$

$$\{\bar{Q}_{\dot\alpha a}, \bar{Q}_{\dot\beta b}\} = 2\sqrt{2}\epsilon_{\dot\alpha\dot\beta}\epsilon_{ab} Z, \tag{5.10}$$

where Z is the $\mathcal{N}=2$ central charge, which commutes with the supersymmetry generators (supercharges) Q^a_α and $\bar{Q}_{\dot\alpha a}$. Due to this property of commutation with the supersymmetry generators, the central charge Z can be fixed to be the eigenvalue of the representation of the algebra.

The composite operators a_α and b_α are given by:

$$a_\alpha = \tfrac{1}{2}[Q^1_\alpha + \epsilon_{\alpha\beta} Q^{\dagger 2}_\beta], \quad b_\alpha = \tfrac{1}{2}[Q^1_\alpha - \epsilon_{\alpha\beta} Q^{\dagger 2}_\beta], \tag{5.11}$$

which permits the re-expression of the $\mathcal{N}=2$ supersymmetry algebra as:

$$\{a_\alpha, a^\dagger_\beta\} = \delta_{\alpha\beta}(M + \sqrt{2}Z), \quad \{b_\alpha, b^\dagger_\beta\} = \delta_{\alpha\beta}(M - \sqrt{2}Z), \tag{5.12}$$

where M is the mass of the $\mathcal{N}=2$ particle states in the representation, and all other anticommutators are equal to zero. All physical states arising from the action of the operators a_α and b_α in the $\mathcal{N}=2$ supersymmetry algebra in Eq. (5.12) on the vacuum state must have positive definite norm. Hence the mass M cannot be less than $\sqrt{2}|Z|$. For $M \neq 0$ this implies the $\mathcal{N}=2$ mass bound $M \geq \sqrt{2}|Z|$. When $M=0$, to satisfy positivity of the norm one must have a trivial central charge, $Z=0$. When the $\mathcal{N}=2$ mass bound is saturated, $M=\sqrt{2}|Z|$, the anticommutator $\{b_\alpha, b^\dagger_\beta\}$ vanishes and the dimension of the representation is decreased. Then $\mathcal{N}=2$ supersymmetry multiplets belonging to the dimensionally reduced representation are referred to as 'short multiplets' (when $M=\sqrt{2}|Z|$). When the mass bound is satisfied by the inequality $M > \sqrt{2}|Z|$, the set of multiplets have representations with unreduced dimensions, they are referred to as a 'long multiplet.'

The $\mathcal{N}=2$ mass bound is identical to the Bogomol'nyi bound, and this coincidence is brought about by the presence of solitonic solutions in the model, namely magnetic monopoles and dyons. The $\mathcal{N}=2$ central charge has a physical origin due to the properties of the $\mathcal{N}=2$ supercharges Q^a_α and $\bar{Q}_{\dot\alpha a}$ [167]. In general the supercharges can be calculated as space integrals of the time component of the supercurrent, which give an expression involving superfields. The anticommutators of the supercharges give rise to non-zero surface terms which cannot be neglected in the presence of electric and magnetic charges. These surface terms then produce the central charge Z. In the absence of matter multiplets, one has:

$$Z = a(q + ig) = ae(n_e + n_m \tau_{cl}), \tag{5.13}$$

where a is the vacuum expectation value of the Higgs scalar field, q and g are respectively the electric and magnetic charges, e is the unit electric charge, n_e and n_m specify the size of the charges, and τ_{cl} is the complexified gauge coupling constant of the classical theory. The value of Z in Eq. (5.13) then implies that $M \geq \sqrt{2}|Z|$ coincides with the BPS mass bound given in Eq. (4.34) in Subsection 4.4.1 of Chapter 4, and can be written explicitly as:

$$M \geq \sqrt{2}|Z| = \sqrt{2}|ae(n_e + n_m \tau_{cl})|. \tag{5.14}$$

Consequently, BPS states, for which $M = \sqrt{2}|Z|$, belong to short multiplets (that is, to reduced representations of the supersymmetry algebra). Furthermore, the mass-charge relation for BPS states is not modified by perturbative or non-perturbative quantum corrections, since it is protected by supersymmetry to all orders. If the mass-charge relation were modified, then the BPS states would no longer satisfy the $\mathcal{N} = 2$ mass bound, and then not belong to a short multiplet, violating supersymmetry. Quantum corrections are also not expected to produce additional degrees of freedom to modify a short multiplet into a long multiplet.

In the low energy effective theory, the mass bound given by Eq. (5.14) is modified and assumes an effective form. The low energy Wilsonian effective action for $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory will be specified by the $\mathcal{N} = 2$ prepotential $\mathcal{F}$, and will be the effective form of the Lagrangian for the $\mathcal{N} = 2$ supersymmetric Yang–Mills theory Eq. (3.86) given in Section 3.4 of Chapter 4. The central charge Z for the effective theory has the form:

$$Z = an_e + a_D n_m, \quad a_D = \frac{\partial \mathcal{F}(a)}{\partial a}, \tag{5.15}$$

where $\mathcal{F}(a)$ is the prepotential in the low energy effective theory given as a function of the classical vacuum expectation value a. We describe the effective action of the $\mathcal{N} = 2$

theory below. In Subsection 5.3.2 below, the central charge in the case when fundamental matter multiplets are present will be described.

R-Symmetry in $\mathcal{N} = 2$ Supersymmetric Yang–Mills Gauge Theory

As was described in Subsection 3.2.3 of Chapter 3, supersymmetric field theories possess an additional symmetry, known as R-symmetry, arising from the supersymmetry algebra, which acts on the supercharges and superfields of the theory. The Lagrangians of $\mathcal{N} = 1$ supersymmetric gauge theories are usually required, for the purposes of renormalizability, to be invariant under a $U(1)$ R-symmetry, as stated in Chapter 3. In the case of $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory, there exists a $U(1)$ R-symmetry and an $SU(2)$ R-symmetry, and these symmetries may be exploited. The former acts upon the $\mathcal{N} = 2$ superspace co-ordinates, and the latter acts upon the indices of the $\mathcal{N} = 2$ supercharges. The $\mathcal{N} = 2$ vector multiplet is comprised of the component fields $\{v_m, \lambda, \psi, \phi\}$. These component fields can be arranged into the $\mathcal{N} = 1$ vector superfield $V(v_m, \lambda)$ and the $\mathcal{N} = 1$ chiral superfield $\Phi(\psi, \phi)$. We denote the $SU(2)$ and $U(1)$ R-symmetries as $SU(2)_R$ and $U(1)_r$, respectively, to distiniguish them from gauge groups and other symmetries. These symmetries act as follows on the component fields of the $\mathcal{N} = 2$ vector multiplet:

$$SU(2)_R \;:\; \lambda \to \lambda' = \phi, \;\; \phi \to \phi' = \lambda, \tag{5.16}$$

$$U(1)_r \;:\; \phi \to \phi' = e^{2i\alpha}\phi, \;\; \psi \to \psi' = e^{i\alpha}\psi, \tag{5.17}$$

$$\lambda \to \lambda' = e^{i\alpha}\lambda, \;\; v_m \to v' = v_m. \tag{5.18}$$

The action of these R-symmetries can be expressed in terms of the $\mathcal{N} = 1$ superfields V and Φ as:

$$U(1)_r \;:\; \Phi(\theta) \to \Phi'(\theta) = e^{2i\alpha}\Phi(e^{-i\alpha}\theta), \;\; V(\theta) \to V'(\theta) = V(e^{-i\alpha}\theta), \tag{5.19}$$

$$U(1)_j \;:\; \Phi(\theta) \to \Phi'(\theta) = \Phi(e^{-i\alpha}\theta), \;\; V(\theta) \to V'(\theta) = V(e^{-i\alpha}\theta), \tag{5.20}$$

where the $U(1)_j$ is a residual symmetry from the action of the $SU(2)_R$ symmetry on the $\mathcal{N} = 2$ component fields. The $U(1)_j$ symmetry is a subgroup of the $SU(2)_R$ symmetry, and is not manifest in the $\mathcal{N} = 1$ superfield decomposition. Hence the $U(1)_j$ symmetry does not rotate the fields λ and ϕ into each other. However, this $U(1)_j \in SU(2)_R$

symmetry can be explicitly expressed as the following transformations on the component fields λ, ϕ and ψ:

$$U(1)_j : \lambda \to \lambda' = e^{i\alpha}\lambda, \quad \phi \to \phi' = e^{i\alpha}\phi, \quad \psi \to \psi' = e^{-i\alpha}\psi. \tag{5.21}$$

The $\mathcal{N} = 2$ scalar multiplet, also known as the $\mathcal{N} = 2$ hypermultiplet, and from which the matter multiplets of the theory can be constructed, has component fields $\{q, \tilde{q}, \chi, \tilde{\chi}\}$ which were introduced in Subsection 3.2.3 of Chapter 3. These fields can be interpreted as being comprised of the two complex scalar fields $\{q, \tilde{q}^\dagger\}$ and the two Weyl spinors $\{\chi, \tilde{\chi}^\dagger\}$, which reside in the fundamental representation of the gauge group. The $SU(2)_R$ and $U(1)_r$ symmetries act as follows on the component fields:

$$SU(2)_R \quad : \quad q \to q' = \tilde{q}^\dagger, \quad \tilde{q}^\dagger \to \tilde{q}'^\dagger = q, \tag{5.22}$$

$$U(1)_r \quad : \quad q \to q' = q, \quad \tilde{q} \to \tilde{q}' = \tilde{q}, \tag{5.23}$$

$$: \quad \chi \to \chi' = e^{-i\alpha}\chi, \quad \tilde{\chi} \to \tilde{\chi}' = e^{-i\alpha}\tilde{\chi}. \tag{5.24}$$

The component fields of the $\mathcal{N} = 2$ scalar multiplet can be arranged into the $\mathcal{N} = 1$ chiral multiplets $Q(q, \chi)$ and $\tilde{Q}(\tilde{q}, \tilde{\chi})$, which are the fundamental chiral matter multiplets familiar from Section 3.3 of Chapter 3. In the $\mathcal{N} = 1$ formulation, the $SU(2)_R$ symmetry is again only manifest as the $U(1)_j$ symmetry. Due to the generic form of the $\mathcal{N} = 1$ superpotential describing the interaction of the vector and scalar multiplets, the $\mathcal{N} =$ 1 superfields Q and $\tilde{Q}$ must have zero R-charge, i.e. they possess neutral R-charge. Therefore the $U(1)_r$ and $U(1)_j$ symmetries act on the scalar multiplets Q and $\tilde{Q}$ as:

$$U(1)_r \quad : \quad Q(\theta) \to Q'(\theta) = Q(e^{-i\alpha}\theta), \quad \tilde{Q}(\theta) \to \tilde{Q}'(\theta) = \tilde{Q}(e^{-i\alpha}\theta), \tag{5.25}$$

$$U(1)_j \quad : \quad Q(\theta) \to Q'(\theta) = e^{i\alpha}Q(e^{-i\alpha}\theta), \quad \tilde{Q}(\theta) \to \tilde{Q}'(\theta) = \tilde{Q}(e^{-i\alpha}\theta). \tag{5.26}$$

As before, the action of the $U(1)_j$ symmetry on the component fields of the $\mathcal{N} = 2$ scalar multiplet can also be expressed explicitly:

$$U(1)_j \quad : \quad q \to q' = e^{i\alpha}q, \quad \tilde{q}^\dagger \to \tilde{q}'^\dagger = e^{-i\alpha}\tilde{q}^\dagger, \tag{5.27}$$

$$: \quad \chi \to \chi' = \chi, \quad \tilde{\chi}^\dagger \to \tilde{\chi}'^\dagger = \tilde{\chi}^\dagger. \tag{5.28}$$

The $U(1)_r$ symmetry is also a chiral symmetry. The classical $\mathcal{N} = 2$ theory is invariant under global $SU(2)_R \times U(1)_r$ transformations. In the quantum $\mathcal{N} = 2$ theory, the $U(1)_r$

symmetry is broken to a discrete subgroup by chiral anomalies. For the pure $\mathcal{N} = 2$ supersymmetric Yang–Mills theory with gauge group $SU(N)$, this discrete subgroup can be determined via instanton methods. In this case the $U(1)_r$ symmetry is broken to a $\mathbb{Z}_{4N}$ discrete symmetry. The discrete symmetry $\mathbb{Z}_{4N}$ can be represented by the generator $e^{2\pi i\alpha}$, where $\alpha = n/4N$, and $n = 1, \ldots, 4N$. The centre of the $SU(2)_R$ symmetry, which acts as $(\lambda, \phi) \to e^{i\pi}(\lambda, \phi)$, is also contained in the $\mathbb{Z}_{4N}$ group, and corresponds to $n = 2N$. Accounting for this leads to the true global R-symmetry group for the quantum $\mathcal{N} = 2$ $SU(N)$ gauge theory, which is given by $SU(2)_R \times \mathbb{Z}_{4N}/\mathbb{Z}_2$. When the Higgs field in the theory acquires a non-zero classical vacuum expectation value, this R-symmetry is broken further. When $N = 2$, which is the case for Seiberg-Witten theory, where the gauge group is $SU(2)$, the R-charge of the Higgs field requires the breaking of the $\mathbb{Z}_{4N}$ symmetry down to a $\mathbb{Z}_4$ symmetry. Hence for the $\mathcal{N} = 2$ $SU(2)$ theory, the complete R-symmetry group is given by $SU(2)_R \times \mathbb{Z}_4/\mathbb{Z}_2$.

Low Energy Effective Action and the $\mathcal{N} = 2$ Prepotential

The form of the low energy Wilsonian effective action for $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory can be deduced implicitly. All of the unknown features of the low energy Wilsonian effective action can be subsumed into the prepotential $\mathcal{F}$, which is a function of massless vector multiplets only, and into higher potential terms which we do not describe here, as they are beyond the scope of this thesis. The $\mathcal{N} = 2$ effective action contains at most two derivatives with respect to the fields of the theory, and is completely specified by the $\mathcal{N} = 2$ prepotential $\mathcal{F}$, whose classical form is known.
In terms of $\mathcal{N} = 1$ superfields, the $\mathcal{N} = 2$ low energy effective action has the form:

$$\mathcal{L}_{\text{eff }\mathcal{N}=2\text{ SYM}} = \frac{1}{8\pi}\left[\int d^4\theta \frac{\partial \mathcal{F}(A)}{\partial A}\bar{A} + \int d^2\theta \frac{1}{2}\frac{\partial^2 \mathcal{F}(A)}{\partial A^2} W^\alpha W_\alpha\right], \tag{5.29}$$

where A is the chiral superfield (formerly written as Φ, which was defined in Eq. (3.31) of Subsection 3.2.3 in Chapter 3) and $\mathcal{F}(A)$ is the prepotential written in terms of A. When the Higgs field, which is the scalar component of A, acquires a non-zero classical vacuum expectation value, which we denote as a, the $SU(2)$ gauge group of the theory is spontaneously broken to the subgroup $U(1)$, and the theory enters the Coulomb phase.

The metric on the space of fields in this case is then given by:

$$ds^2 = g_{ij} da^i d\bar{a}^j = \mathrm{Im}\,(\tau) da d\bar{a}, \tag{5.30}$$

where the effective prepotential $\mathcal{F}(a)$ enters as:

$$\tau(a) = \frac{\partial^2 \mathcal{F}(a)}{\partial a^2}, \tag{5.31}$$

in which τ is the effective complexified coupling constant, given by:

$$\tau(a) = \tau_{\text{eff}}(a) = \frac{4\pi i}{g^2(a)} + \frac{\vartheta}{2\pi}. \tag{5.32}$$

The classical prepotential in the low energy effective theory is known and can be obtained by using the procedure to obtain the Wilsonian effective action. The classical prepotential $\mathcal{F}(A)$ is given by:

$$\mathcal{F}_{\text{cl}}(A) = \tfrac{1}{2}\tau_{\text{cl}} A^2, \tag{5.33}$$

where τ_{cl} is the classical complexified gauge coupling constant:

$$\tau_{\text{cl}} = \frac{4\pi i}{g^2} + \frac{\vartheta}{2\pi}. \tag{5.34}$$

in which g is the classical coupling constant, which does not run (i.e. it does not depend upon a). This uses the definition of the complexified gauge coupling constant in Eq. (2.17) in Section 2.2 of Chapter 2, but the classical and quantum forms of τ must be distinguished between in this chapter.

The perturbative form of $\mathcal{F}(A)$ can be determined following the methods introduced by Seiberg for the exact construction of superpotentials as described in Section 4.2 of Chapter 4 [169]. The microscopic $\mathcal{N} = 2$ supersymmetric $SU(2)$ theory is specified by the Lagrangian in Eq. (3.86) with gauge group $SU(2)$ in Section 3.4 of Chapter 3. This is an asymptotically free theory, and so valid perturbative calculations can be performed at high energies within it. The $U(1)_r$ symmetry of the classical theory is broken by the standard chiral anomaly in the quantum theory. This anomaly is manifest as the chiral current J_5^m given by:

$$\partial_m J_5^m = -\frac{N}{16\pi^2} v_{mn}{}^* v^{mn}. \tag{5.35}$$

This result implies that to one loop order in perturbation theory the low energy effective Lagrangian changes under a $U(1)_r$ transformation by an amount:

$$\delta \mathcal{L}_{\text{eff } \mathcal{N}=2 \text{ SYM}} = -\frac{\alpha N}{16\pi^2} v_{mn}{}^* v^{mn}, \tag{5.36}$$

where α is the phase of the $U(1)_r$ transformation.

This shift in the effective Lagrangian translates the ϑ-angle. This translation can be used to set $\vartheta = 0$ by a chiral rotation of the fermion fields in the Lagrangian. Hereon we adopt the value $\vartheta = 0$ for the vacuum angle.

By requiring that $\mathcal{L}_{\text{eff } \mathcal{N}=2 \text{ SYM}}$ changes by the amount in Eq. (5.36) under a $U(1)_r$ transformation, the one loop perturbative contribution to $\mathcal{F}$ can be determined. Given the form of the variation in Eq. (5.36), only terms quadratic in the gauge field strength v_{mn} are relevant as it is from these terms which the variation must originate. Therefore the terms which contribute to the variation are:

$$\begin{aligned}\frac{1}{16\pi}\mathrm{Im}\left[\mathcal{F}''(e^{2i\alpha}A)(-v_{mn}v^{mn} + iv_{mn}{}^{*}v^{mn})\right] \\ = \frac{1}{16\pi}\mathrm{Im}\left[\mathcal{F}''(A)(-v_{mn}v^{mn} + iv_{mn}{}^{*}v^{mn})\right] - \frac{\alpha N}{8\pi^2}v_{mn}{}^{*}v^{mn}, \quad (5.37)\end{aligned}$$

in which we have cancelled a factor of two. The prepotential is then required to satisfy:

$$\mathcal{F}''(e^{2i\alpha}A) = \mathcal{F}''(A) - \frac{2\alpha N}{\pi}. \quad (5.38)$$

This condition can be re-expressed as a differential equation for $\mathcal{F}(A)$ when α is infinitesimal:

$$\frac{\partial^3\mathcal{F}(A)}{\partial A^3} = \frac{N}{\pi}\frac{i}{A}. \quad (5.39)$$

Integrating this condition yields the one loop perturbative correction to $\mathcal{F}(A)$:

$$\mathcal{F}_{\text{1-loop}}(A) = \frac{i}{2\pi}A^2\ln\frac{A^2}{\Lambda^2}, \quad (5.40)$$

where Λ is a fixed dynamically generated scale which acts as an energy cutoff. The one loop contribution to the prepotential is the only perturbative quantum corrections which $\mathcal{F}(A)$ receives. This is due to $\mathcal{N} = 2$ supersymmetry and is related to the one loop perturbative exactness of the $\mathcal{N} = 2$ beta function. This is also consistent with the one loop nature of the chiral anomaly in Eq. (5.35) and the absence of infra-red divergences in the theory.

Although the $\mathcal{N} = 2$ prepotential is perturbatively exact to one loop, it still receives non-perturbative quantum corrections, as Seiberg has argued [169]. These non-perturbative quantum corrections arise from instanton effects in the theory. The prepotential receives only instanton effects and not anti-instanton effects because the prepotential $\mathcal{F}(A)$ is a holomorphic function [169]. A k-instanton correction to $\mathcal{F}(A)$ arising from a k-instanton configuration will be proportional to the k-instanton action $\exp(-8\pi^2 k/g^2)$. The one loop exact Wilsonian beta function of the theory, given by Eq. (5.5), then implies that such a k-instanton factor can be written as:

$$e^{-8\pi^2 k/g^2} = \left(\frac{\Lambda}{A}\right)^{4k}, \tag{5.41}$$

which arises from a relation between the classical instanton action and the one loop exact Wilsonian beta function in Eq. (5.5) [169]. Seiberg noted that the broken $U(1)_r$ symmetry of the theory is restored if the scale Λ is assigned an R-charge of 2. If this is done, the prepotential $\mathcal{F}(A)$ must possess an R-charge of 4. Since the chiral superfield A has an R-charge of 2, one then expects the k-instanton correction to be proportional to A^2. Combining the classical prepotential $\mathcal{F}_{\text{cl}}(A)$, Eq. (5.33), with the perturbative prepotential $\mathcal{F}_{\text{1-loop}}(A)$, Eq (5.40), and the generic non-perturbative contribution, the form of the low energy effective prepotential $\mathcal{F}(A)$ is then:

$$\begin{aligned} \mathcal{F}(A) &= \mathcal{F}_{\text{cl}}(A) + \mathcal{F}_{\text{1-loop}}(A) + \mathcal{F}_{k\text{-instanton}}(A) \\ &= \tfrac{1}{2}\tau_{\text{cl}} A^2 + \frac{i}{2\pi} A^2 \ln \frac{A^2}{\Lambda^2} + \sum_{k=1}^{\infty} \mathcal{F}_k \left(\frac{\Lambda}{A}\right)^{4k} A^2, \end{aligned} \tag{5.42}$$

where the coefficients $\mathcal{F}_k$ are independent of the fields in the theory and are numerical constants. These coefficients are independent of the fields because instantons contribute to the path integral only via zero modes in supersymmetric gauge field theories. We will return to this aspect of the prepotential in Chapter 6. Seiberg and Witten were able to propose an exact form for $\mathcal{F}(A)$ which includes a determination of all of the coefficients $\mathcal{F}_k$, thus completely specifying the form of the prepotential.

We note that, in the limit of a large vacuum expectation value a in the quantum theory, the theory becomes weakly coupled via asymptotic freedom. In this limit the quantum theory can be approximated by replacing the classical coupling g with the running coupling $g(a)$ and so $g(a)$ and $\mathcal{F}(a)$ can be determined from perturbation theory by integrating the one loop beta function.

Parameterization of the $\mathcal{N} = 2$ Moduli Space

We first describe the moduli space of classical $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory in the Coulomb phase. In the classical theory, which is given by the microscopic Lagrangian Eq. (3.86) of Section 3.4 of Chapter 3 with gauge group $SU(2)$, there exists the scalar potential V, where:

$$V(\phi) = -\frac{1}{2g^2}\mathrm{Tr}\left([\phi^\dagger, \phi]^2\right), \tag{5.43}$$

which was given in Eq. (3.83) of Section 3.4 in Chapter 3. The field ϕ is the scalar field component of the chiral superfield A, which is identified with the Higgs field. The Higgs vacuum is defined by $[\phi^\dagger, \phi] = 0$. This definition implies that ϕ assumes values in the Cartan subalgebra of the gauge group G of the theory, so that $\phi = \phi_i H^i$, where H is the subgroup of G generated by the elements of the Cartan subalgebra. The Higgs field may also be in the trivial Higgs vacuum, for which $\phi = 0$. The gauge group G is generically broken to the subgroup H in the Higgs vacuum. Elements of the coset G/H are gauge transformations which connect physically equivalent vacua, and the Higgs vacuum is thus not invariant under these transformations. However, physically inequivalent vacua are specified by the different vacuum expectation values which ϕ can assume, which are denoted by ϕ_i. Hence the degrees of freedom represented by ϕ_i parameterize the space of physically inequivalent vacua, or the moduli space, of the theory. The dimension of the moduli space is equal to the rank r of the gauge group G. There remains a residual gauge invariance within this parameterization arising from the elements of the coset G/H. Transformations under these coset elements do not leave the Higgs vacuum invariant but do leave ϕ as an element of the Cartan subalgebra. Such transformations correspond to Weyl reflections of the subgroup H. The gauge invariant parameterization of the moduli space will thus be Weyl invariant functions of ϕ. General formulae for the Weyl invariants of a group can be used to explicitly construct such functions.

For the gauge group $SU(2)$, the classical Higgs field is given by

$$\phi = \tfrac{1}{2}a\sigma_3, \tag{5.44}$$

where σ_3 is the third standard Pauli spin matrix. The Weyl invariant for this classical field is given by u_{cl}, where

$$u_{\mathrm{cl}} = \mathrm{Tr}(\phi^2) = \tfrac{1}{2}a^2. \tag{5.45}$$

The Weyl invariant quantity $u_{\rm cl}$ may be referred to as the classical modulus for the $SU(2)$ theory. In the case of the quantum $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, the quantum moduli space will be parameterized by the vacuum expectation values of the classical Weyl invariant. We refer to this Weyl invariant quantity as the quantum modulus $u_{\rm qu}$, which is given by the quantum vacuum expectation value:

$$u_{\rm qu} = \langle \mathrm{Tr}(\phi^2) \rangle. \tag{5.46}$$

In the classical limit, $u_{\rm qu}$ shall tend to its classical value, $u_{\rm cl}$:

$$\lim_{\hbar \to 0} u_{\rm qu} \to u_{\rm cl} = \tfrac{1}{2} a^2.$$

Seiberg and Witten were able to determine $u_{\rm qu}$ for the low energy effective $SU(2)$ theory in terms of the classical vacuum expectation value a.

Electric-Magnetic Duality in $\mathcal{N} = 2$ Supersymmetric Yang–Mills Gauge Theory

Given the one loop contribution to the $\mathcal{N} = 2$ prepotential $\mathcal{F}(A)$, the requirement of a positive definite kinetic energy for the theory reveals that the theory does not possess a single global description. Using the one loop contribution in Eq. (5.40) and the definition of $\tau(a)$ given in Eq. (5.31), one can deduce that for large $|a|$, the complexified gauge coupling constant becomes:

$$\tau(a) = \frac{i}{\pi}\left(\ln\left(\frac{a^2}{\Lambda^2}\right) + 3\right). \tag{5.48}$$

The form of $\tau(a)$ in Eq. (5.48) is not a single valued function, whereas the the field space metric given by $\mathrm{Im}\,\tau$ in Eq. (5.30) is single valued. The imaginary part of $\tau(a)$, $\mathrm{Im}\,\tau$, is a harmonic function and so does not possess a global minimum. It can possess local minima, but if it is a globally defined function, local minima can only occur if $\tau(a)$ is negative in some regions. If $\tau(a)$ were positive everywhere, then a global minimum could be defined, which would contradict the harmonicity of $\mathrm{Im}\,\tau$. To maintain positive kinetic energy, $\tau(a)$ can only be defined locally, and so admits no global description. Thus the moduli space of the theory has no global description. In the regions where $\tau(a)$ becomes negative, a different description of the theory must be employed which is appropriate locally, as a theory in which a global $\tau(a)$ can assume both positive and negative values is unphysical. A candidate alternative description of any gauge theory is its electric-magnetic dual, or more generally, its S-dual. The concepts of electric-magnetic duality and S-duality were introduced in Chapter 4 in the context of $\mathcal{N}=4$ supersymmetric Yang–Mills gauge theory. In Chapter 4 we also described the notions of phases and the moduli space of vacua for $\mathcal{N}=1$ supersymmetric Yang–Mills gauge theories.

The dual description of the theory can be reached by appropriate duality transformations. These transformation can be deduced from the gauge field terms of the action, which appear in the bosonic part of the action. In the Abelian low energy theory, in four dimensional Euclidean spacetime, one has $(v_{mn})^2 = (^*v_{mn})^2$ and $^*(^*v_{mn}) = v_{mn}$ for the Abelian gauge field strength v_{mn}, and so the gauge field terms have the following schematic form:

$$\frac{1}{32\pi}\mathrm{Im}\int \tau(a)(v_{mn}+i^*v_{mn})^2 \;=\; \frac{1}{16\pi}\mathrm{Im}\int \tau(a)(v_{mn}v^{mn}+i^*v_{mn}v^{mn}). \tag{5.49}$$

A Lagrange multiplier V_D, which is also a vector superfield, may be used to implement the Bianchi identity $\epsilon^{mnkl}\partial_m v^{mn} \equiv 0$. We use a normalization for V_D such that all fundamental $SU(2)$ matter fields have half integer charges. This convention implies that a magnetic monopole of the theory obeys $\epsilon^{0nkl}\partial_n v_{kl} = 8\pi\delta^3(x)$. By coupling the multiplier V_D to such a monopole, the Lagrange multipler term is schematically given by:

$$\frac{1}{8\pi}\int V_{Dm}\epsilon^{mnkl}\partial_n v_{kl} \;=\; \frac{1}{8\pi}\int {}^*v_{Dmn}v^{mn} \;=\; \frac{1}{16\pi}\mathrm{Re}\int(^*v_{Dmn}-iv_{Dmn})(v^{mn}+i^*v^{mn}), \tag{5.50}$$

in which implements the Bianchi identity for v_{mn} aforementioned, and where $v_{Dmn} = \partial_m V_{Dn} - \partial_n V_{Dm}$ is the dual gauge field strength in the sense of duality ($v_{Dmn} \neq^* v_{mn}$). The Lagrange multipler term Eq. (5.50) can be added to the gauge field terms in Eq. (5.49). Integrating the resulting terms with respect to v_{mn} yields the dual gauge field terms:

$$\frac{1}{32\pi}\mathrm{Im}\int\left(-\frac{1}{\tau(a)}\right)(v_{Dmn}+i^*v_{Dmn})^2 \;=\; \frac{1}{16\pi}\mathrm{Im}\int\left(-\frac{1}{\tau(a)}\right)(v_{Dmn}v^{Dmn}+i^*v_{Dmn}v^{Dmn}). \tag{5.51}$$

Hence the duality transformations in the $\mathcal{N}=2$ theory induce the gauge field v_m which couples to electric charges to be replaced with the dual gauge field v_{Dm} which couples to magnetic charges. It also transforms the complexified gauge coupling to:

$$\tau(a) \rightarrow \tau'(a) = \tau_D(a) = -\frac{1}{\tau(a)}, \tag{5.52}$$

which is of the same form as the electric-magnetic duality described in Chapter 4. Hence the candidate duality of the theory is indeed realized as a form of electric-magnetic duality. Furthermore, the gauge field term Eq. (5.49) is invariant under the translation:

$$\tau \rightarrow \tau' = \tau + 1. \tag{5.53}$$

The transformations given in Eqs. (5.52,5.53) result from the action of the following matrices on the space of scalar fields:

$$S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{5.54}$$

Together, the matrices in Eq. (5.54) generate the duality, or modular, group $SL(2,\mathbb{Z})$. Thus the theory exhibits a form of S-duality, which was described in Chapter 4. The full $SL(2,\mathbb{Z})$ duality transformation acts upon $\tau(a)$ as:

$$\tau(a) \rightarrow \tau'(a) = \frac{\alpha\tau(a)+\beta}{\gamma\tau(a)+\delta}, \tag{5.55}$$

where $\alpha\delta - \beta\gamma = 1$ and $\{\alpha, \beta, \gamma, \delta\} \in \mathbb{Z}$. The dual gauge field description of the theory is related to a dual scalar field description of the theory by $\mathcal{N} = 2$ supersymmetry. The function of scalar fields $\{A\}$ defined by $h(A) = \partial\mathcal{F}/\partial A$ is related to a function of the dual scalar fields $\{A_D\}$ denoted by $h_D(A_D)$, and the exact relation between them can be found by applying a method similar to that for the gauge field terms above to the analagous scalar field terms. Then the duality transformations imply that $h = A_D$ and $h_D = -A$, and again the duality group generated by the transformations is the S-duality group $SL(2,\mathbb{Z})$.

Mathematically, the S-duality of the gauge and scalar fields in the theory can be made manifest in the metric on the space of fields as given in Eq. (5.30). The field space metric is invariant under the S-duality group $SL(2,\mathbb{Z})$ when expressed in terms of the dual scalar field component (classical vacuum expectation value) a_D, which is the (electric-magnetic) dual of a:

$$ds^2 = \mathrm{Im}\,(\tau)\, da_D d\bar{a} = -\frac{i}{2}(da_D d\bar{a} - da d\bar{a}_D), \tag{5.56}$$

where ds^2 is the metric on the moduli space.

The classical moduli space of the $SU(2)$ theory is a manifold of one complex dimension, which we denote $\mathcal{M}_{SU(2)}$. If u is a holomorphic co-ordinate on $\mathcal{M}_{SU(2)}$, the metric Eq. (5.56) can be written in terms of u. This will be of use later as the co-ordinate u can be identified with the moduli space parameter $\langle \mathrm{Tr}\phi^2 \rangle$. The functions $\{a(u), a_D(u)\}$ can be interpreted as parameters on a complex space isomorphic to $\mathbb{C}^2$ which define an $SL(2,\mathbb{Z})$ bundle over $\mathcal{M}_{SU(2)}$. The metric of the moduli space can be written in terms of u as:

$$ds^2 = \mathrm{Im}\frac{da_D}{du}\frac{d\bar{a}}{d\bar{u}}dud\bar{u} = -\frac{i}{2}\left(\frac{da_D}{du}\frac{d\bar{a}}{d\bar{u}} - \frac{d\bar{a}_D}{d\bar{u}}\frac{da}{du}\right)dud\bar{u}. \tag{5.57}$$

This expression of the metric is manifestly $SL(2,\mathbb{Z})$ invariant, and the previous form of the metric, Eq. (5.56), is recovered when $u = a$ is set. For arbitary functions $\{a(u), a_D(u)\}$, the metric Eq. (5.57) can be negative. However, positivity of the kinetic term in the low energy effective action requires that the metric ds^2 also be positive. Thus in the low energy effective theory, the functions $\{a(u), a_D(u)\}$ cannot be arbitary, and must be such that the moduli space metric is positive.

Dyons, Coupling and Monodromy of the $\mathcal{N} = 2$ Moduli Space

We first describe the dyon spectrum of the $\mathcal{N} = 2$ $SU(2)$ theory, which will be used later in the determination of the $\mathcal{N} = 2$ moduli space. In the microscopic $SU(2)$ theory, the BPS mass bound is given by $M \geq \sqrt{2}|Z|$, where Z is the central charge given by:

$$Z = a(n_e + \tau_{cl} n_m), \tag{5.58}$$

in which τ_{cl} is the classical complexified gauge coupling defined in Eq. (5.34). In the low energy effective theory, Z becomes modified. When there is a non-zero vacuum expectation value, $a \neq 0$, any matter field multiplets in the theory will become massive. The form of the central charge for matter states of the theory is fixed by $\mathcal{N} = 2$ supersymmetry. If a matter multiplet possesses an electric charge n_e, then its contribution to the central charge will be $Z_e = an_e$. Matter fields which possess magnetic charge n_m, such as magnetic monopoles, contribute an amount $Z_m = a_D n_m$ to the central charge. Hence the general central charge in the $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills theory will be:

$$Z = an_e + a_D n_m, \tag{5.59}$$

which determines the mass spectrum for dyons of charges (n_e, n_m) via the BPS mass bound.

If the moduli space of the theory, $\mathcal{M}_{SU(2)}$, is taken to possess non-trivial structure, which may include, for instance, singularities, then the vector $v^T = (a_D, a)$, taken about a closed loop, will be transformed under the action of the monodromy group of the moduli space. That is, the vector taken about the loop will be transformed to a different vector due to properties of the moduli space present in the vicinity of the vector.

The mass formula for dyons given by:

$$M \geq \sqrt{2}|Z| = \sqrt{2}|an_e + a_D n_m| \tag{5.60}$$

must be invariant under monodromy transformations of the moduli space because the mass is a physically observable quantity. If the vector $v = (a, a_D)$ is transformed to $\mathrm{M}v$, where $\mathrm{M} \in Sp(4, \mathbb{Z})$ is the monodromy matrix, then the vector $w = (n_m, n_e)$ will be transformed under the action of the monodromy group as $w\mathrm{M}^{-1}$. In a similar way, the monodromy acts on the complexified gauge coupling as:

$$\tau \to \tau' = \frac{A\tau + B}{C\tau + D}, \quad \mathrm{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{5.61}$$

where M is the monodromy matrix. This transformation is isomorphic to that for the transformation of the period matrix of a genus r Riemann surface under the monodromy group of the moduli space of genus r Riemann surfaces. Then the quantum moduli space of vacua, $\mathcal{M}_{SU(2)}$, could be identified with the moduli space of a genus one Riemann surface, $\mathcal{M}_{g=1}$. More precisely, these moduli spaces could be conjectured to be isomorphic: $\mathcal{M}_{SU(2)} \simeq \mathcal{M}_{g=2}$. A Riemann surface is a connected complex analytic manifold of one complex dimension. The most simple Riemann surface is the complex plane, $\mathbb{C}$, which is a non-compact Riemann surface. We refer the reader to the standard text on Riemann surfaces for a detailed exposition of the mathematics of Riemann surfaces [249]. Assuming this conjecture to be true, the variables a and a_D can be calculated from the periods of the Riemann surface. From the variables a and a_D, the prepotential of the low energy effective action can be calculated. It is this hypothesis which forms part of the core of Seiberg-Witten theory, and from which the proposed low energy effective action is derived. Later in this subsection the evidence put forward by Seiberg and Witten in support of this identification will be described. This will involve determining the monodromies of the moduli space using physical and mathematical intuition.

The Beta Function of $U(1)$ Theories

To assist with the determination of the monodromies on the moduli space, we now describe the beta function for a fundamental matter multiplet interacting with a $U(1)$ gauge theory. From the $\mathcal{N}=2$ supersymmetric matter multiplet, there exist, in terms of $\mathcal{N}=1$ superfields, Weyl spinors, complex scalars and chiral superfields, assuming the matter multiplet is a reduced (short) $\mathcal{N}=2$ multiplet of spin less than or equal to $\frac{1}{2}$. Then Weyl spinors and complex scalars, of equal electric charge Q, contribute to the beta function of the $U(1)$ gauge theory by an amount given by:

$$\beta_{matter}(g) \equiv \mu\frac{dg}{d\mu} = \frac{g^3}{8\pi^2}Q^2. \tag{5.62}$$

The equation defining the beta function, Eq. (5.62), can be re-written in terms of the normalized gauge coupling constant $\alpha_g = g^2/4\pi$ and the coefficient of the factor g^3, which we denote b_0. Then the beta function is defined by:

$$\mu\frac{d}{d\mu}\left(\frac{1}{\alpha_g}\right) = -8\pi b_0 = -\frac{1}{\pi}Q^2. \tag{5.63}$$

With the ϑ-angle set to zero via a chiral rotation of the fermion states of the theory, as described earlier in this subsection, the complexified gauge coupling constant $\tau(a)$ then becomes $\tau(a) = i/\alpha_g$. Using this, the re-expressed beta function Eq. (5.63) can be written in terms of $\tau(a)$ as:

$$\mu\frac{d\tau}{d\mu} = -\frac{i}{\pi}Q^2. \tag{5.64}$$

The scale μ of the theory can be naturally identified with the vacuum expectation value a. Choosing the matter multiplet charge to be $Q=1$, Eq. (5.64) then yields $\tau(a)$ as:

$$\tau(a) \simeq -\frac{i}{\pi}\ln\left(\frac{a}{\Lambda}\right), \tag{5.65}$$

where Λ is the dynamically generated scale of the theory. This will assist in the determination of the monodromies of the moduli space at finite u, where u is the holomorphic co-ordinate on the quantum moduli space, $\mathcal{M}_{SU(2)}$, of the theory.

This calculation can be performed for the case of a magnetic multiplet, that is, a multiplet of magnetic monopoles. The contribution of this multiplet to the beta function of a $U(1)$ gauge theory can be calculated in a similar way using the dual variables of the theory, giving an expression for the dual complexified gauge coupling $\tau_D(a_D)$. If the magnetic charge of the multiplet is also set equal to unity, then the dual of Eq. (5.65) is given by:

$$\tau_D(a_D) \simeq -\frac{i}{\pi}\ln\left(\frac{a_D}{\Lambda}\right). \tag{5.66}$$

The dual formula Eq. (5.66) will not be used in the remainder of this subsection, but we include it here for completeness.

The Structure of the $\mathcal{N}=2$ Moduli Space

We now turn to the determination of the monodromies of the $\mathcal{N}=2$ moduli space. Knowledge of the moduli space, and particularly the monodromy structure of the moduli space, will be useful in deriving the form of the prepotential for the low energy effective $\mathcal{N}=2$ theory.

The variable used to specify regions of the $\mathcal{N}=2$ moduli space will be the holomorphic co-ordinate u of the moduli space. We denote this holomorphic co-ordinate generically as u, which shall be equal to either u_{cl} or u_{qu} depending on whether the classical or quantum moduli space of vacua is under consideration. For the case of $\mathcal{N}=2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, the moduli space of vacua has the form of a complex plane, and so we refer to the generic state of the moduli space of vacua as the u-plane. When $|a|$ is large, u is also large, the theory is asymptotically free and $u \to u_{cl} = \frac{1}{2}a^2$. The prepotential can be approximated in this regime by:

$$\mathcal{F}(a) = \frac{i}{2\pi}a^2\ln\left(\frac{a^2}{\Lambda^2}\right). \tag{5.67}$$

Using this limit of the prepotential, the dual vacuum expectation value for large $|a|$ is then:

$$a_D = \frac{\partial\mathcal{F}}{\partial a} = \frac{2ia}{\pi}\ln\left(\frac{a}{\Lambda}\right) + \frac{ia}{\pi}. \tag{5.68}$$

A closed loop (isomorphic to a circle) on the u-plane about the point $u = 0$, that is, a monodromy transformation about the singularity $u = 0$, induces the following transformations on $\ln u$ and $\ln a$:

$$\ln u \rightarrow \ln u + 2\pi i, \tag{5.69}$$

$$\ln a \rightarrow \ln a + i\pi, \tag{5.70}$$

which corresponds to changes in the variables a_D and a by amounts, respectively, of:

$$a_D \rightarrow -a_D + 2a, \tag{5.71}$$

$$a \rightarrow -a. \tag{5.72}$$

Hence the action of the monodromy group at large u is produced by the monodromy matrix M_∞, given by:

$$M_\infty = PT^{-4} = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix}, \tag{5.73}$$

where P is the negative of the identity element of the $SL(2, \mathbb{Z})$ group, given by:

$$P = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.74}$$

and T is a matrix defined in Eq. (5.54) which acts upon the vector $(a_D, a)^T$. The monodromy M_∞ acts upon the magnetic and electric quantum numbers of the BPS states as $(n_m, n_e) \rightarrow (-n_m, -n_e - 2n_m)$, a transformation which leaves the BPS mass bound invariant, in accordance with the physical observability of mass.

The monodromy at large u implies that there exist other monodromies elsewhere on the u-plane. The reason for this is as follows. The monodromy group will be Abelian if there exist other monodromy matrices which commute with M_∞, or if there exist no other monodromies. An Abelian monodromy group would imply that the variable a is a valid global co-ordinate, which contradicts the need for local descriptions of the theory, and thus violates positivity of the kinetic energy. Therefore the monodromy group must be non-Abelian, and there exist other monodromies on the u-plane which are different from the 2×2 unit matrix. To obtain a non-Abelian monodromy group, there must exist at least two singularities in the u-plane at finite u with non-trivial monodromy. The moduli space at finite u corresponds to the non-perturbative regime of the theory. The singularities on the u-plane will be related via the discrete ($\mathbb{Z}_2$) symmetry $u \rightarrow -u$. A closed loop containing both of these singularities should yield a monodromy matrix equal to M_∞.

The singularities on the moduli space must have a physical origin. Seiberg and Witten intuit that these singularities arise from the low energy Wilsonian effective action. The Wilsonian effective action is obtained by integrating out all of the massive states and modes above a chosen energy scale. The massive particle states will produce a non-trivial metric on the moduli space through their monodromy loops. The masses of such states will depend on the co-ordinate u. However, the masses of some of these states may vanish for specific values of u. All of the massive states and modes have been integrated out of the action, and only massless (or 'light') states (and modes) remain in it. When these massive states dependent on u become massless, singularities appear at these points on the moduli space. This is because they have been integrated out of the action, when massless states (and modes) are to be retained in the procedure to obtain the Wilsonian effective action, without divergences occurring due to massless states becoming massive. The form of these singularities on the moduli space, and hence the monodromies associated with them, are dependent on the properties of the massive particle states which exhibit this behaviour. At finite u, the mass spectrum is not known in terms of u, and so the determination of these additional singularities must be acheived indirectly. Some assumptions are necessary to obtain further information about the theory at this stage. Given the reflection symmetry which is known to relate the other singularities on the moduli space, Seiberg and Witten assume that some generic particle states become massless at the points $u = 1$ and $u = -1$, which is the simplest choice for these singular points. Let the monodromies associated with these singularities be M_1 and M_{-1}, respectively. As aforementioned, it is anticipated that these matrices satisfy the condition $M_1 M_{-1} = M_\infty$ from which the massless states may be extracted.

The massive states in the theory with spin less than or equal to 1 consists of the gauge boson multiplet, comprised of gauge fields, and the dual gauge boson multiplet, comprised of monopoles and dyons. The gauge boson multiplet becomes massless in the classical theory at $u = 0$. This value of u may become non-zero in the quantum theory. Seiberg and Witten argue that a massive spin 1 multiplet becoming massless in the theory is inconsistent. Therefore the singularities must arise when the dual gauge boson multiplet becomes massless. This multiplet has a spin of no greater than $\frac{1}{2}$ and therefore resides in a short representation of the $\mathcal{N} = 2$ supersymmetry algebra. This implies that the particle states in this multiplet are BPS states, and hence are non-perturbative states. To calculate the monodromies of the massless states in this multiplet is the next step towards finding the structure of the moduli space. However, the monopole and dyon multiplet cannot be locally coupled to the fundamental matter fields of the theory. Instead, dual gauge fields can be coupled to monopoles and dyons in a local manner exactly analogous to the local coupling between gauge fields and electrically charged states. Therefore the electric-magnetic dual description of the theory, as described above, may be used.

Then, to determine the monodromy for a generic massless monopole or dyon state in the theory, only the monodromy of the singularity arising when a massive electrically charged multiplet becomes massless is required. An electric-magnetic duality transformation of the resulting monodromy will give the monodromy for the dual (that is, magnetic) singularity. Using the description of the theory dual to the magnetic theory, the monodromies of the massless dual gauge fields may then be calculated, hence giving information about the singularities on $\mathcal{M}_{SU(2)}$

In the dual description of the electric theory, the monopoles and dyons will appear as elementary states. Let the magnetic theory be described by the dual variables a^d and a_D^d, where the superscript d denotes 'dual.' In the vicinity of a generic massive monopole or dyon state becoming massless, all massive states of the theory can still be integrated out of the action and the remaining theory is a $U(1)$ gauge theory coupled to a fundamental matter multiplet. The vacuum expectation value of the scalar field in the dual description is a^d. A massive BPS state of unit electric charge becomes massless at the point $a^d = 0$, which occurs on the moduli space at the point which we label $u = u^d$. In the vicinity of the point $u = u^d$, the variable a^d is a valid local co-ordinate, which may be expanded as $a^d \approx c^d(u - u^d)$. The one loop beta function for the $U(1)$ theory coupled to a fundamental matter multiplet was described previously in this subsection, and may be employed here. Near the point $u = u^d$, the one loop beta function for this theory will have the following form in the dual description:

$$\tau(a^d) = -\frac{i}{\pi}\ln\left(\frac{a^d}{\Lambda}\right). \tag{5.75}$$

From this form of $\tau(a^d)$, the dual variable a_D^d is found to be:

$$a_D^d = -\frac{i}{\pi}a^d\ln\left(\frac{a^d}{\Lambda}\right) + \frac{i}{\pi}. \tag{5.76}$$

A closed loop on the u-plane about u^d (for which $(u - u^d) \to e^{2\pi i}(u - u^d)$) will induce the following monodromy:

$$a_D^d \to a_D^d + 2a^d, \tag{5.77}$$

$$a^d \to a^d. \tag{5.78}$$

To find the monodromy of the singularity appearing when a dyon of charge (n_m, n_e) becomes massless, Seiberg and Witten begin by assuming that a dyon of charge $(0, 1)$ will be an elementary state in the dual (magnetic) description of the theory. A generic $SL(2, \mathbb{Z})$ duality transformation acting on the dual description vectors $(a_D^d, a^d)^T$ and $(n_m^d, n_e^d)^T$ transforms these to:

$$\begin{pmatrix} a_D^d \\ a^d \end{pmatrix} \to \begin{pmatrix} \alpha a_D + \beta a \\ \gamma a_D + \delta a \end{pmatrix}, \tag{5.79}$$

$$\begin{pmatrix} n_m^d \\ n_e^d \end{pmatrix} \to \begin{pmatrix} \gamma a_m - \delta n_e \\ -\beta n^m + \alpha n_e \end{pmatrix}, \tag{5.80}$$

where $\alpha\delta - \beta\gamma = 1$, and under which the central charge Z defined in Eq. (5.59) is invariant. We now set $n_m^d = 0$ and $n_e^d = 1$. Then the variables a_D^d and a^d describe the dyon coupling to the dual gauge field in precisely the same way in which the variables a_D and a describe unit electric charges coupled to gauge fields. The monodromy of the singularity produced when the $(0, 1)$ charge dyon becomes massless is given by Eqs. (5.77,5.78). To deduce the action of the monodromy Eqs. (5.77,5.78) on the original variables a_D and a, one can invert Eq. (5.79) to obtain these variables in terms of the dual description variables a_D^d and a^d as:

$$a_D = -\beta a^d + n_e a_D^d, \tag{5.81}$$

$$a = \alpha a^d - n_m a_D^d. \tag{5.82}$$

Using Eq. (5.80), we can deduce the remaining unknowns α and β in terms of the quantum numbers (n_m, n_e). Then the action of the monodromy given in Eqs. (5.77,5.78) on the vector $(a_D, a)^T$ is:

$$\begin{pmatrix} a_D \\ a \end{pmatrix} \rightarrow \begin{pmatrix} 1+2n_e n_m & 2n_e^2 \\ -2n_m^2 & 1-2n_e n_m \end{pmatrix} \begin{pmatrix} a_D \\ a \end{pmatrix} = M(n_m, n_e) \begin{pmatrix} a_D \\ a \end{pmatrix}, \qquad (5.83)$$

where we have denoted the generic dyon monodromy matrix as $M(n_m, n_e)$.

The monodromies at the points $u = \pm 1$ can now be calculated. Seiberg and Witten assume that a charge (m, n) dyon becomes massless at $u = 1$ and that a charge (m', n') dyon becomes massless at $u = -1$. The monodromies of these states must then obey the following condition:

$$M_1(m,n)M_{-1}(m',n') = M_\infty. \qquad (5.84)$$

Given M_∞ in Eq. (5.73), this can be expressed using the generic dyon monodromy $M(m, n)$ of Eq. (5.83) as:

$$\begin{pmatrix} 1+2mn & 2n^2 \\ -2m^2 & 1-2mn \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -0 & -1 \end{pmatrix} \begin{pmatrix} 1-2m'n' & -2n'^2 \\ 2m'^2 & 1+2m'n' \end{pmatrix}. \qquad (5.85)$$

The component equations of the matrix equation Eq. (5.85) are solved by the values $m = \pm 1$ and $m' = \pm 1$. These values of m and m' can be used to determine n' in terms of n, for which the following possible solution sets exist:

$$(m,n) \ : \ (1,n) \ , \ (-1,n) \ , \ (-1,n) \ , \ (1,n), \qquad (5.86)$$

$$(m',n') \ : \ (1,n-1) \ , \ (1,-n-1) \ , \ (-1,n+1) \ , \ (-1,-n+1). \qquad (5.87)$$

These solutions indicate that only dyons of unit magnetic charge may contribute to the monodromy. The semiclassical result that only such dyons are stable is consistent with this result.

The most simple solution of Eq. (5.85) is that for which $m = m' = 1$, $n = 0$ and $n' = -1$. These values give the monodromy matrices of the dyon singularities as:

$$M_1 = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \quad M_{-1} = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix}, \qquad (5.88)$$

which satisfy the requirement that $M_1 M_{-1} = M_\infty$.

To interpret these singularities physically, one can consider the action of the monodromy.

Under a monodromy transformation, in general, the quantum numbers of the dyon will change. The dyons which become massless and which result in a singularity will remain invariant under the action of the monodromy. This is because these particular dyons give rise to the monodromy matrices. For a generic dyon of charge (q_m, q_e), the eigenvalue equation for the action of the monodromy matrix $M(m,n)$ is:

$$(q_m, q_e)\begin{pmatrix} 1+2mn & 2n^2 \\ -2m^2 & 1-2mn \end{pmatrix} = (q_m, q_e). \tag{5.89}$$

This eigenvalue equation implies that $nq_m - mq_e = 0$, assuming that m and n are both non-zero. This condition is satisfied by $q_m = m$ and $q_e = n$. If the dyon is to be stable, then this is the unique solution of the eigenvalue equation Eq. (5.89). Therefore, if the monodromy matrix for a singularity is known, the dyon which produces the singularity can be deduced.

For the monodromy matrices M_1 and M_{-1}, the eigenvalue equation Eq. (5.89) implies that: the monodromy matrix M_1 results from a dyon of charge $(n_m, n_e) = (1,0)$, that is, a magnetic monopole, becoming massless at $u = 1$; the monodromy matrix M_{-1} results from a dyon of charge $(n_m, n_e) = (1,-1)$ becoming massless at $u = -1$. Then at the point where the monopole becomes massless, one has $a_D = 0$, and at the point where the dyon becomes massless, one has $a = a_D$, due to the vanishing of the mass bound (and thus $|Z|$) at the values of (n_m, n_e) given. The monodromy at M_∞ is interpreted as arising from an electric photon, for which $(n_m, n_e) = (0,0)$ and the mass bound vanishes automatically. Furthermore, at the singularity at infinity, M_∞ shifts the electric charge by two units (by $2n_e$). Therefore, at the singularities of the u-plane, the electric charge is only defined modulo 2, and so the absolute value of the electric charge is not known.

The Seiberg–Witten Elliptic Curve

Now that the monodromies of the moduli space of vacua $\mathcal{M}_{SU(2)}$ have been determined, the conjectured identification of $\mathcal{M}_{SU(2)}$ with the moduli space of a genus one Riemann surface $\mathcal{M}_{g=1}$ will enable the calculation of the variables a and a_D in terms of the holomorphic co-ordinate u.

This is essentially the Riemann–Hilbert problem of reconstructing a holomorphic function from knowledge of its singularities, that is, from its monodromies [249].

The moduli space $\mathcal{M}_{SU(2)}$ is the u-plane with singularities at the points $u = 1$, $u = -1$ and the asymptote $u \to \infty$. The monodromies for these singular points are, respectively, M_1, M_{-1} and M_∞, which obey $M_1 M_{-1} = M_\infty$. At finite u, points on the u-plane are related by the $\mathbb{Z}_2$ symmetry $u \to -u$. There exists a flat $SL(2,\mathbb{Z})$ bundle on $\mathcal{M}_{SU(2)}$ which has the vector $(a_D, a)^T$ as a holomorphic section. The variables to be determined exhibit the following asymptotic behaviour:

$$u \to \infty \ : \ \begin{cases} a \ \approx \ \sqrt{2u} \\ a_D \ \approx \ i(\sqrt{2u}/\pi)\ln u \end{cases}, \tag{5.90}$$

$$u = 1 \ : \ \begin{cases} a \ \approx \ c_0(u-1) \\ a_D \ \approx \ a_0 + (i/\pi) a_D \ln a_D \end{cases}, \tag{5.91}$$

$$u = -1 \ : \ \begin{cases} a \ \approx \ c_0(u-1) \\ (a - a_D) \ \approx \ a_0 + (i/\pi)(a - a_D)\ln(a - a_D) \end{cases}, \tag{5.92}$$

where c_0 and a_0 are constants. The metric on the moduli space is given by $ds^2 = \mathrm{Im}(\tau)|da|^2$, where the complexified gauge coupling constant can be expressed in terms of a and a_D as:

$$\tau(u) \ = \ \frac{da_D(u)/du}{da(u)/du}. \tag{5.93}$$

For the kinetic energy of the theory to remain positive, the quantity $\mathrm{Im}(\tau)$ must be positive definite. The three monodromies of $\mathcal{M}_{SU(2)}$ generate the subgroup $\Gamma(2)$ of $SL(2,\mathbb{Z})$. The moduli space $\mathcal{M}_{SU(2)}$ can be described as the quotient of the upper half complex plane H by $\Gamma(2)$, which corresponds to a Riemann surface of genus one. Then the three singularities of the u-plane (i.e. the moduli space) are the three cusp points of this quotient.

The u-plane $H/\Gamma(2)$, which is a compact Riemann surface, can be described by a family of elliptic curves E_u. This is due to the Torelli theorem in the study of Riemann surfaces and algebraic curves [250]. The Torelli theorem implies that all of the information about a compact Riemann surface is contained in its (normalized) period matrix [250]. More precisely, the Torelli theorem states the necessary and sufficient condition under which two compact Riemann surfaces are isomorphic; this occurs when their normalized period matrices are equivalent, under a suitable canonical homology basis. The period matrix shall depend upon differential forms which can be written in terms of an algebraic curve. The theorem that any compact Riemann surface of genus g can be represented as the normalization of a plane algebraic hyperelliptic curve of degree $(2g+2)$ then enables one to always specify a Riemann surface via an algebraic (hyper-)elliptic curve [250].

In this case, the set of elliptic curves E_u which represent the u-plane $H/\Gamma(2)$ already exist in the relevant mathematics literature, and are specified by:

$$y^2 = (x-1)(x+1)(x-u), \tag{5.94}$$

where y and x are complex dummy variables. The elliptic curve in Eq. (5.94) is referred to as the 'Seiberg–Witten elliptic curve,' as it is the elliptic curve proposed to describe the moduli space of vacua in Seiberg–Witten theory. This elliptic curve is invariant under the discrete transformations $u \to -u, x \to -x, y \to \pm iy$, which together generate a $\mathbb{Z}_4$ symmetry, of which the subgroup $\mathbb{Z}_2$ acts on the co-ordinate u. These symmetries are also present on $\mathcal{M}_{SU(2)}$. The x-plane (given by the real and complex parts of x) has a topology constrained by the requirement that y^2 is a single valued function. This requires that the x-plane be a double cover of the complex plane $\mathbb{C}$ with the point at infinity added to it. The x-plane also possesses four branch points, at $\{-1, 1, u, \infty\}$, two pairs of which are joined by cuts. This space is topologically indistinguishable from a genus one Riemann surface, which is a complex torus, in which the cuts correspond to cycles on the Riemann surface [249]. A loop which intersects both cuts on the x-plane corresponds to the other cycle on the Riemann surface. To calculate a and a_D on the Riemann surface, a homology basis is required. Let γ_1 and γ_2 be two independent homology cycles normalized such that they have unit intersection number:

$$\gamma_1 \circ \gamma_2 = 1. \tag{5.95}$$

The cycles γ_1 and γ_2 are homology one-cycles because they form a local basis for the first homology group $H^1(E_u, \mathbb{C})$ for the set of curves E_u. The one-cycles γ_1 and γ_2 also vary continuously with u. Each homology one-cycle γ_i can be associated (or paired) with an element λ_i of the first cohomology group $H_1(E_u, \mathbb{C})$ for E_u:

$$\gamma_i \to b_{ij} = \oint_{\gamma_i} \lambda_j, \tag{5.96}$$

where $i, j = 1, 2$ and the elements λ_i can be interpreted as meromorphic one-forms on E_u. These meromorphic one-forms have vanishing residue modulo exact differential forms [249]. The pairing between λ_i and γ_i is invariant under continuous deformations of γ_i at all points of λ_i, including poles, due to the vanishing residues of λ_i. Through this association with the cycles γ_i, the λ_i are also elements of $H_1(E_u, \mathbb{C})$.

Seiberg and Witten choose the following basis for the one-forms on E_u:

$$\lambda_1 = \frac{dx}{y}, \quad \lambda_2 = \frac{xdx}{y}, \tag{5.97}$$

where, up to a multiplicative (scalar) constant, λ_1 is the unique holomorphic differential form on E_u. Using the definition for b_{ij} in Eq. (5.96), the complex torus can be characterised by the parameter τ_u, given by:

$$\tau_u = \frac{b_{11}}{b_{21}}, \tag{5.98}$$

such that $\mathrm{Im}(\tau_u) > 0$. Let λ be an arbitrary section of $H^1(E_u, \mathbb{C})$; that is, let:

$$\lambda = a_1(u)\lambda_1 + a_2(u)\lambda_2. \tag{5.99}$$

Then, to relate the variables a and a_D to the formalism describing the genus one Riemann surface, Seiberg and Witten assume that a and a_D are given by:

$$a_D = \oint_{\gamma_1} \lambda, \quad a = \oint_{\gamma_2} \lambda. \tag{5.100}$$

The system of integral equations in Eq. (5.100) can be solved using the Picard–Fuchs equations [193], but here we follow the method of evaluation of Seiberg and Witten. To be consistent with the symmetries of the BPS mass bound with no matter terms, the differential form λ must have only vanishing residues at its poles. This then implies that on circling a singularity, a and a_D transform in accordance with the transformation properties of γ_1 and γ_2 under a subgroup of $SL(2, \mathbb{Z})$. Both a and a_D then transform purely under $SL(2, \mathbb{Z})$.

The arbitrariness in λ can be fixed via the positivity condition $\mathrm{Im}(\tau) > 0$ for $\mathcal{M}_{SU(2)}$. Seiberg and Witten propose that λ is related to λ_1 via the relation:

$$\frac{d\lambda}{du} = f(u)\lambda_1 = f(u)\frac{dx}{y}. \tag{5.101}$$

Differentiating Eq. (5.100) with respect to u and using Eq. (5.101) then gives:

$$\frac{da_D}{du} = f(u)b_{11}, \quad \frac{da}{du} = f(u)b_{21}. \tag{5.102}$$

Comparing these expressions to Eq. (5.98) for τ_u yields:

$$\tau = \frac{b_{11}}{b_{21}} = \frac{da_D/du}{da/du} = \tau_u. \tag{5.103}$$

Since $\tau = \tau_u$, and $\mathrm{Im}(\tau_u) > 0$, therefore $\mathrm{Im}(\tau) > 0$, and positivity of the kinetic energy is ensured. Seiberg and Witten argue that the converse of this argument is true;, that is, starting from τ, one would conclude that $\tau_u = \tau$ is true. Thus $d\lambda/du$ is independent of λ_2. The asymptotic behaviour of the theory near the singularities of the u-plane imply that the function $f(u)$ has the form $f(u) = -\sqrt{2}/4\pi$. The differential form λ is then given by:

$$\lambda = \frac{\sqrt{2}}{2\pi}\frac{dx\sqrt{x-u}}{\sqrt{x^2-1}} = \frac{\sqrt{2}}{2\pi}\frac{\sqrt{y}dx}{x^2-1} = \frac{\sqrt{2}}{2\pi}\frac{dx}{\sqrt{y}}(x-u). \tag{5.104}$$

To explicitly calculate a and a_D it is necessary to select a particular basis of homology cycles on E_u. Let γ_2 be the a-cycle of the complex torus, which is the loop about the points $(-1, 1)$ on the x-plane. This loop can be deformed so that it becomes as close as possible to a line between the points $x = -1$ and $x = 1$. Then using Eq. (5.100), $a(u)$ is given by:

$$a(u) = \frac{\sqrt{2}}{\pi}\int_{-1}^{1}\frac{dx\sqrt{x-u}}{\sqrt{x^2-1}}. \tag{5.105}$$

With this choice for the one-cycle γ_2, accordingly γ_1 can be chosen to be the b-cycle of the complex torus, which is the loop about the points $(1, u)$ on the x-plane. Again using Eq. (5.100), this gives $a_D(u)$ as:

$$a_D(u) = \frac{\sqrt{2}}{\pi}\int_{1}^{u}\frac{dx\sqrt{x-u}}{\sqrt{x^2-1}}. \tag{5.106}$$

The expressions Eq. (5.105) and Eq. (5.106), for $a(u)$ and $a_D(u)$, respectively, can be expanded in x and their asymptotic behaviour at the singularities of the u-plane checked. These expressions exhibit the required behaviour near the singularities, which validates the choice of one-cycles made by Seiberg and Witten.

Determination of the $\mathcal{N} = 2$ Prepotential

The integrals Eq. (5.105) and Eq. (5.106) defining $a(u)$ and $a_D(u)$ can be identified with particular hypergeometric functions in their integral representation. In turn, these hypergeometric functions can be expressed as complete elliptic integrals. The integral representation of a general hypergeometric function $F(\alpha,\beta,\gamma;z)$ is given by:

$$\begin{aligned} F(\alpha,\beta,\gamma;z) &= \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)}\int_0^1 dx\, x^{\beta-1}(1-x)^{\gamma-\beta-1}(1-zx)^{-\alpha} \\ &= \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\beta)}\sum_{n\geq 0}\frac{\Gamma(\alpha+n)\Gamma(\beta+n)}{\Gamma(\gamma+n)}\frac{z^n}{n!}. \end{aligned} \tag{5.107}$$

Using this representation of the hypergeometric function $F(\alpha,\beta,\gamma;z)$, Eq. (5.105) implies that a(u) may be written as:

$$a(u) = \sqrt{2(1+u)}F\left(-\tfrac{1}{2},\tfrac{1}{2},1;\tfrac{2}{1+u}\right). \tag{5.108}$$

To assist in identifying Eq. (5.106) and hence $a_D(u)$ with a hypergeometric function, the substitution $x=(u-1)t+1$ in Eq. (5.106) gives:

$$a_D(u) = \frac{i}{\pi}(u-1)\int_0^1 dt\, t^{-\frac{1}{2}}(1-t)^{\frac{1}{2}}\left(1-\tfrac{(1-u)t}{2}\right)^{-\frac{1}{2}}. \tag{5.109}$$

Comparing Eq. (5.109) with Eq. (5.107), $a_D(u)$ can be expressed in terms of hypergeometric functions as:

$$a_D(u) = \frac{i}{2}(u-1)F\left(\tfrac{1}{2},\tfrac{1}{2},2;\tfrac{(1-u)}{2}\right). \tag{5.110}$$

Furthermore, making use of the complete elliptic integrals defined by:

$$K\left(\frac{\sqrt{2}}{\sqrt{(1+u)}}\right) = \frac{\pi}{2}F\left(\tfrac{1}{2},\tfrac{1}{2},1;\tfrac{2}{(1+u)}\right), \tag{5.111}$$

$$E\left(\frac{\sqrt{2}}{\sqrt{(1+u)}}\right) = \frac{\pi}{2}F\left(-\tfrac{1}{2},\tfrac{1}{2},1;\tfrac{2}{(1+u)}\right), \tag{5.112}$$

enables the variables $a(u)$ and $a_D(u)$ to be expressed as in terms of these as:

$$a(u) = \frac{4\sqrt{1+u}}{\sqrt{2}\pi}E\left(\frac{\sqrt{2}}{\sqrt{(1+u)}}\right), \tag{5.113}$$

$$a_D(u) = \frac{4\sqrt{1+u}}{\sqrt{2}\pi i}\left(E\left(\sqrt{1-\tfrac{2}{(1+u)}}\right)-K\left(\sqrt{1-\tfrac{2}{(1+u)}}\right)\right). \tag{5.114}$$

The complexified gauge coupling can now be calculated in terms of u. The result is:

$$\tau(u) = \frac{da_D/du}{da/du} = \frac{iK\left(\sqrt{1-\frac{2}{(1+u)}}\right)}{K\left(\frac{\sqrt{2}}{\sqrt{(1+u)}}\right)}. \tag{5.115}$$

The explicit solution for the prepotential of the low energy Wilsonian effective action of four dimensional $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory follows from substituting Eq. (5.115) into Eq. (5.31) above and integrating over the scalar vacuum expectation values $a(u)$. The problem of determining the prepotential from the system of equations above can also be formulated as a set of Picard–Fuchs equations [193], which provide an alternative method of obtaining the same solution. The expressions for $a(u)$ and $a_D(u)$ can be expanded in terms of u to check their behaviour under monodromy and also their asymptotic limits.

Monopole Condensation and Confinement in $\mathcal{N} = 2$ Supersymmetric Yang-Mills Gauge Theory

We now describe confinement in the Seiberg–Witten solution for the low energy Wilsonian effective action of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang-Mills gauge theory. The phenomenon of confinement is described in terms of the $\mathcal{N} = 2$ theory broken to an $\mathcal{N} = 1$ theory by the addition of a mass term for the chiral superfield multiplet Φ. The microscopic Lagrangian also has a mass gap, the existence of which is required for confining behaviour. This assumes the form of an $\mathcal{N} = 1$ superpotential $\mathcal{W}$, given by:

$$\mathcal{W} = m\mathrm{Tr}(\Phi^2), \tag{5.116}$$

where m is the mass of the chiral superfield Φ. The mass m is a bare mass. The low energy $\mathcal{N} = 1$ theory resulting from the modified $\mathcal{N} = 2$ theory is an Abelian $U(1)$ gauge theory which possesses a $\mathbb{Z}_4$ chiral symmetry. Let the term $\mathrm{Tr}(\Phi^2)$ be represented by a chiral superfield U in the low energy theory. Then the scalar component of U is the holomorphic function $u_{\mathrm{qu}} = \langle \mathrm{Tr}(\phi^2)\rangle$, which parameterizes the quantum moduli space, as has already been described. When m is small, the form of the effective superpotential is mU, which is added to the low energy effective Lagrangian. This modification is expected to remove the vacuum degeneracy of the theory and give mass to the scalar multiplet of component fields U. If the theory is to have a mass gap, then the $U(1)$ gauge fields are required to be massive also. This can be done by adding low mass (or 'light') gauge fields to the theory, or by a Higgs mechanism involving the existing charged low mass gauge fields. If additional low mass gauge fields were added, this would result in a strongly coupled non-Abelian gauge theory.

However, additional low mass fields cannot occur in the theory because this would give rise to extra singularities on the moduli space $\mathcal{M}_{SU(2)}$. As reasoned by Seiberg and Witten, and described above, there should exist only three singularities on $\mathcal{M}_{SU(2)}$, all of which are accounted for in their analysis.

Thus, in order to obtain a mass gap in the theory, the existing charged low mass gauge fields should give rise to a Higgs mechanism. Via electric-magnetic duality, this can be described in terms of dual (charged) gauge fields, that is, magnetic monopole and dyons. The complete $\mathcal{N} = 1$ superpotential can be written in the dual description as:

$$\mathcal{W} = \sqrt{2} A_D M \tilde{M} + mU(A_D), \tag{5.117}$$

where M, $\tilde{M}$ are $\mathcal{N} = 1$ chiral superfields which comprise the $\mathcal{N} = 1$ magnetic monopole and dyon multiplet, and A_D is the dual of the scalar field A. The low energy vacua of the theory then occur when the variation of the superpotential is zero, $d\mathcal{W} = 0$, and $|M| = |\tilde{M}|$. When $m = 0$, then $M = \tilde{M} = 0$ in the vacua of the theory. Then a_D is arbitrary and the $\mathcal{N} = 2$ moduli space $\mathcal{M}_{SU(2)}$ is recovered, since the vacuum is unmodified. When $m \neq 0$, the variation of $\mathcal{W}$ results in:

$$\sqrt{2} M \tilde{M} + m \frac{du}{dA_D} = 0, \tag{5.118}$$

$$a_D M = a_D \tilde{M} = 0. \tag{5.119}$$

If $du/dA_D \neq 0$, then both M and $\tilde{M}$ are non-zero. Then $a_D = 0$ and one obtains:

$$M = \tilde{M} = -\left(\frac{m}{\sqrt{2}}\right)^{1/2} \frac{du(0)}{dA_D}. \tag{5.120}$$

The superfields M and $\tilde{M}$ are charged, and according to Eq. (5.120), their vacuum expection values are non-zero. Via the Higgs mechanism, these superfields then generate a mass for the low mass gauge fields. In this way, the mass gap of the microscopic theory is reproduced. As the multiplet $M\tilde{M}$ is magnetically charged, the dual theory has a magnetic Higgs mechanism. This occurs when $M = \tilde{M} \neq 0$, and through the Higgs mechanism, massless magnetic monopoles condense in the vacuum. Electric charges are then confined in the electric theory by the dual (magnetic) Meissner effect through electric-magnetic duality. The rôle of magnetic monopoles in the confinment mechanism of Yang–Mills gauge theories has thus been given an exact, if conjectural, realization.

In this subsection we have briefly outlined how Seiberg and Witten were able to deduce the low energy Wilsonian effective action of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory by indirect methods. The concepts of electric-magnetic duality, $\mathcal{N} = 2$ supersymmetry, moduli space and holomorphy have been employed alongside powerful techniques of complex analysis to propose the exact form of this low energy effective theory. The physical consequences of this include a proposed explanation and derivation of the mass gap and the confinement of electric charges in the microscopic theory.
In the Subsection 5.3.2 below we briefly outline the methods Seiberg and Witten used to extend their work on pure $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills theory to $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD, which is the pure Yang–Mills theory coupled to fundamental matter multiplets.

5.3.2 $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD

We now describe the exact determination of the low energy Wilsonian effective action for $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with N_f fundamental matter multiplets. The analysis of Seiberg and Witten [171] for this case is lengthy and varies for differing values of N_f. Many elements of the analysis for the pure $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills theory described in the Subsection 5.3.1 are used.
The theory under consideration is given by the ($\mathcal{N} = 1$ superfield) Lagrangian in Eq. (3.87) with gauge group $SU(2)$ of Section 3.4 in Chapter 3. The Wilsonian beta function for $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with N_f matter multiplets is given by:

$$\beta(g) = -\frac{(4 - N_f)g^3}{16\pi^2}, \tag{5.121}$$

using the result for the $\mathcal{N} = 2$ Wilsonian beta function in Eq. (5.5) of Section 5.2. In order for the theory to remain asymptotically free, one must have $N_f \leq 4$ (for $N_f \geq 0$). Theories with this range of N_f shall only be described in this subsection.
Following Seiberg and Witten, a different charge normalization is chosen for the analysis

of this theory. This is to maintain integral electric charges for the particle states in the theory when fundamental matter multiplets are present. In the pure $\mathcal{N} = 2$ supersymmetric Yang–Mills theory, all fields transform in the adjoint representation and all particle states possess integer charges, given by the quantum numbers n_e and n_m appearing in the central charge Z. When fundamental matter is added to this theory, n_e may assume half integer values. To maintain integer values of n_e in $\mathcal{N} = 2$ supersymmetric QCD, one can multiply n_e by 2 and divide a by 2, so that the mass bound is unchanged. We summarize these changes in conventions as:

$$n_e \rightarrow 2n_e, \quad a \rightarrow \tfrac{1}{2}a, \quad Z \rightarrow Z. \tag{5.122}$$

The dual vacuum expectation value a_D is left unchanged, but is now defined according to an equation different to Eq. (5.68) of Subsection 5.3.1. The dual variable a_D is now given by:

$$a_D = \frac{1}{2}\frac{\partial \mathcal{F}}{\partial a}. \tag{5.123}$$

The asymptotic behaviour of both a and a_D is also modified due to the different charge normalization. When the holomorphic co-ordinate u becomes large, $|u| \rightarrow \infty$, the variables a and a_D tend to:

$$a \approx \tfrac{1}{2}\sqrt{2u}, \quad a_D = \frac{4i}{\pi} a \log a. \tag{5.124}$$

Given these rescalings of a and a_D, the complexified gauge coupling constant is modified accordingly, and assumes the form:

$$\tau = \frac{\partial a_D}{\partial a} = \frac{\vartheta}{\pi} + \frac{8\pi i}{g^2}. \tag{5.125}$$

The form of the elliptic curves which are proposed to provide the low energy effective action of the theory are also modified by this change in the charge normalization, as will be described below. The physical content of the theory will remain unchanged by this alteration in convention.

Seiberg and Witten consider the case of $\mathcal{N} = 2$ supersymmetric QED as preparation for applying the same techniques used for the $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory to the $\mathcal{N} = 2$ supersymmetric QCD. This is because in the Coulomb phase of these theories, the $SU(2)$ gauge symmetry is broken to a $U(1)$ symmetry. Furthermore, in the Higgs phase of $\mathcal{N} = 2$ supersymmetric QED, the metric on the moduli space of the theory can be determined uniquely by the symmetries of the theory and receives no quantum perturbative or non-perturbative corrections. However, we will proceed to describe the full $\mathcal{N}=2$ supersymmetric QCD theory, beginning with the central charge and special symmetries of the theory.

Central Charge, Special Symmetries and Phases of $\mathcal{N}=2$ Supersymmetric $SU(2)$ QCD

The central charge for the $\mathcal{N}=2$ supersymmetric $SU(2)$ Yang–Mills gauge theory is changed when fundamental matter multiplets are added to the Lagrangian. The masses of the matter fields will also contribute to the central charge Z. In $\mathcal{N}=1$ notation, the matter multiplets are composed of the chiral superfields Q and $\tilde{Q}$. These have $\mathcal{N}=1$ component fields (q,χ) and $(\tilde{q},\tilde{\chi})$, respectively. These transform in the $\mathbf{2}$ and $\bar{\mathbf{2}}$ representations of $SU(2)$. The component fields belong to a multiplet with spin less than or equal to $\frac{1}{2}$, and so reside in a reduced representation of the $\mathcal{N}=2$ supersymmetry algebra. The saturated mass bound $M=\sqrt{2}|Z|$ is expected to hold, but it does so only if the contributions of the bare masses of the multiplet are included.

The fundamental matter multiplet gives rise to an $\mathcal{N}=1$ superpotential in the Lagrangian for $\mathcal{N}=2$ supersymmetric $SU(2)$ QCD. This is in addition to the kinetic and gauge coupling terms which will involve Q and $\tilde{Q}$. For N_f fundamental matter multiplets in the theory, the $\mathcal{N}=1$ superpotential $\mathcal{W}$ has the form:

$$\mathcal{W} = \sum_{i=1}^{N_f}\left[\sqrt{2}\tilde{Q}_i\Phi Q_i + m_i\tilde{Q}_iQ_i + \sqrt{2}Q_i^\dagger\Phi^\dagger\tilde{Q}_i^\dagger + m_iQ_i^\dagger\tilde{Q}_i^\dagger\right], \tag{5.126}$$

where Φ is the chiral superfield of the (adjoint) vector multiplet, and m_i are the masses of the i matter multiplets, where $i=1,\ldots,N_f$. The first term and third term are associated with coupling the matter fields to the $\mathcal{N}=2$ fields, and the second and fourth terms are $\mathcal{N}=2$ supersymmetry invariant mass terms for the matter fields. If all masses m_i are equal, then the theory also possesses an $SU(N_f)$ 'flavour' symmetry, in analogy to ordinary QCD. If some of the masses m_i are unequal, the $SU(N_f)$ symmetry is broken to smaller subgroups. When all masses m_i are unequal, the $SU(N_f)$ symmetry is completely broken to $U(1)^{N_f}$.

A calculation of the contribution of the superfields Q and $\tilde{Q}$ to the supercurrents of the theory results in an extra term which corresponds to the $U(1)$ subgroup factors of the broken $SU(N_f)$ symmetry. This extra term modifies the central charge Z of the theory accordingly, with the result that Z has the form:

$$Z = an_e + a_D n_m + \sum_i^{N_f} \frac{1}{\sqrt{2}} m_i S_i, \tag{5.127}$$

where the S_i are the space integrals given by:

$$S_i = \int d^3x \left(D_0 q_i^\dagger q_i + q_i D_0 q_i^\dagger - \tfrac{i}{2}\chi_i^\dagger \chi_i + \tfrac{i}{2}\chi_i \chi_i^\dagger + D_0 \tilde{q}_i^\dagger \tilde{q}_i + \tilde{q}_i D_0 \tilde{q}_i^\dagger - \tfrac{i}{2}\tilde{\chi}_i^\dagger \tilde{\chi}_i + \tfrac{i}{2}\tilde{\chi}_i \tilde{\chi}_i^\dagger \right), \tag{5.128}$$

which occur in the anti-commutators of the supercharges when matter multiplets are taken into account.

A further special case occurs when all masses are zero, $m_i = 0$. The theory then possesses a global $SU(N_f) \times SU(2)_R \times U(1)_r$ symmetry. For the particular gauge group $SU(2)$, this 'flavour' symmetry is enhanced to the larger symmetry of $O(2N_f)$. This originates from the property that for the gauge group $SU(2)$, the fundamental and anti-fundamental representations are isomorphic, so that the Q^i and $\tilde{Q}_i$ matter multiplets can be arranged into vectors of dimension $2N_f$ invariant under $O(2N_f)$. Additionally, when the gauge group is $SU(2)$, there is a parity symmetry associated with this. This parity symmetry, denoted ρ, is a $\mathbb{Z}_2 \subset O(2N_f)$ symmetry which acts as:

$$\rho \; : \; Q_1 \leftrightarrow \tilde{Q}_1, \tag{5.129}$$

leaving all other fields invariant. This parity transformation is used later in the analysis of the theory. Due to the presence of a $\mathbb{Z}_2$ symmetry in the R-symmetry $\mathbb{Z}_2 \subset U(1)$, the global symmetry group of the theory is $O(2N_f) \times SU(2)_R \times U(1)_r/\mathbb{Z}_2$.

We now outline the phases of the theory. The phase the theory is in depends on the number of fundamental matter multiplets N_f. When $N_f = 0$ and $N_f = 1$, the theory exists only in the Coulomb phase. In this phase the vacuum expectation values are non-zero, $\langle \phi \rangle \neq 0$, and the $SU(2)$ gauge symmetry is broken to a $U(1)$ subgroup. For $N_f = 1$, the matter multiplets acquire masses through the Higgs mechanism. The $U(1)_r$ symmetry is spontaneously broken due to the R-charge of the superfield Φ.

When $N_f \geq 2$, the theory can exist in either a Coulomb or Higgs phase. In the Higgs phase, the gauge symmetry is completely broken and there are no electric or magnetic charges present. In this thesis we do not describe this phase further, and describe only work concerning the Coulomb phase.

R-Symmetry in $\mathcal{N}=2$ Supersymmetric $SU(2)$ QCD

As described in Subsection 5.3.1, there exist R-symmetries in supersymmetric gauge theories. For $\mathcal{N}=2$ supersymmetric $SU(2)$ QCD, the theory possesses an $SU(2)_R$ symmetry, from which $U(1)_r$ and $U(1)_j$ symmetries arise. The $U(1)_r$ symmetry is generically broken to the discrete symmetry $\mathbb{Z}_{4N-2N_f}$ due to instanton effects in correlation functions. When the gauge group is $SU(2)$ (that is, $N=2$) there exists an additional $O(2N_f)$ symmetry, which is realized as the parity transformation ρ given in Eq. (5.129). This symmetry group is anomalous because it changes the sign of correlation functions in the theory. Since the theory is invariant under this parity transformation, $U(1)_r$ transformations are permitted which also change the sign of correlation functions, but must transform only this. To correct the sign of these functions, and eliminate the freedom to perform such transformations, an anomalous $\mathbb{Z}_2$ transformation can be introduced. Therefore, for $N=2$, the $U(1)_r$ symmetry is broken to the discrete subgroup $\mathbb{Z}_{4(4-N_f)}$. When combined with the newly introduced anomalous $\mathbb{Z}_2$ symmetry, the $\mathbb{Z}_{4(4-N_f)}$ symmetry is then anomaly free and acts as follows on the superfields of the theory:

$$\begin{aligned}
W_\alpha &\rightarrow W'_\alpha = \omega W_\alpha(\omega^{-1}\theta), \\
\Phi &\rightarrow \Phi' = \omega^2 \Phi(\omega^{-1}\theta), \\
Q^1 &\rightarrow Q^{1\prime} = \tilde{Q}_1(\omega^{-1}\theta), \\
\tilde{Q}_1 &\rightarrow \tilde{Q}'_1 = Q^1(\omega^{-1}\theta), \\
Q^i &\rightarrow Q^{i\prime} = Q^i(\omega^{-1}\theta), i \neq 1 \\
Q_i &\rightarrow Q_{i\prime} = \tilde{Q}_i(\omega^{-1}\theta), i \neq 1,
\end{aligned} \tag{5.130}$$

where $\omega = \exp(2\pi i/4(4-N_f))$. When $N_f = 0$, there are no matter fields to cancel the anomaly, with the consequence that only the square of the above transformations are anomaly free, thus reproducing Eqs. (5.19,5.19,5.25,5.26) in Subsection 5.3.1. The transformations in Eq. (5.130) can be combined with the transformations of the $U(1)_j$

symmetry subgroup (recall that $U(1)_j \subset SU(2)_R$). The result is a $\mathbb{Z}_{4(4-N_f)}$ symmetry which commutes with the $\mathcal{N} = 1$ supersymmetry algebra. This symmetry acts on the superfields of the theory as:

$$\begin{aligned}
\Phi &\to \Phi' = \omega^2\Phi(\theta), \\
Q^1 &\to Q^{1\prime} = \omega^{-1}\tilde{Q}_1(\theta), \\
\tilde{Q}_1 &\to \tilde{Q}_1' = \omega^{-1}Q^1(\theta), \\
Q^i &\to Q^{i\prime} = \omega^{-1}Q^i(\theta), i \neq 1 \\
Q_i &\to Q_{i\prime} = \omega^{-1}\tilde{Q}_i(\theta), i \neq 1.
\end{aligned} \tag{5.131}$$

The $\mathbb{Z}_{4(4-N_f)}$ symmetry is broken to a $\mathbb{Z}_4$ symmetry due to the transformation properties of the holomorphic co-ordinate (also known as the order parameter) u on the quantum moduli space. Under the $\mathbb{Z}_{4(4-N_f)}$ transformations, u transforms as:

$$u \to u' = \exp\left[\frac{2\pi i}{(4-N_f)}\right] u. \tag{5.132}$$

The resulting $\mathbb{Z}_4$ symmetry acts non-trivially on the u-plane.
The R-symmetry can be used to fix the generic form of the variables a and a_D in the theory. The arguments which lead to these were originally made by Seiberg. At large $|u|$, the perturbative behaviour of $a(u)$ and $a_D(u)$ is determined by the one loop exact beta function Eq. (5.121), and is given by:

$$a(u) \approx \tfrac{1}{2}\sqrt{2u} + \cdots , \tag{5.133}$$

$$a_D(u) \approx \frac{(4-N_f)}{i}2\pi a(u)\ln\left(\frac{u}{\Lambda^2_{N_f}}\right) + \cdots , \tag{5.134}$$

where the dots represent non-perturbative corrections due to instantons, and $\Lambda^2_{2N_f}$ is the dynamically generated scale associated with the theory. This is analogous to the case when $N_f = 0$, which was described in Subsection 5.3.1. A generic k-instanton correction will be proportional to the k-instanton factor given by:

$$e^{-8\pi^2 k/g^2} = \left(\frac{\Lambda_{N_f}}{a}\right)^{k(4-N_f)}. \tag{5.135}$$

The broken $U(1)_r \times \mathbb{Z}_2$ symmetry can be restored by assigning u and Λ_{N_f} charges under the $U(1)_r$ and $\mathbb{Z}_2$ transformations. For the theory to be invariant under these transformations, a factor of $\sqrt{u}$ is required before the k-instanton corrections. Furthermore, the metric on the moduli space (that is, the u-plane), is invariant under the $\mathbb{Z}_2$ parity transformations ρ. This implies that only k-instanton configurations with k equal to an odd integer do not contribute to a and a_D. The generic forms of a and a_D are then:

$$a(u) = \tfrac{1}{2}\sqrt{2u}\left[1 + \sum_{n=1}^{\infty} a_n \left(\frac{\Lambda_{N_f}^2}{u}\right)^{n(4-N_f)}\right], \qquad (5.136)$$

$$a_D(u) = \frac{2\pi(4-N_f)}{i} a(u) \ln\left(\frac{u}{\Lambda_{N_f}^2}\right) + \sqrt{u}\sum_{n=1}^{\infty} a_{Dn}\left(\frac{\Lambda_{N_f}^2}{u}\right)^{n(4-N_f)}, \qquad (5.137)$$

where $\{a_n\}$ and $\{a_{Dn}\}$ are numerical coefficients. The theory has massive charged states in its Coulomb phase and the remaining unbroken global symmetry acts on these states.

BPS States in $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD

In the Coulomb phase of the theory BPS states can occur. There are two types of BPS states present. Firstly there are saturated BPS states arising from the matter multiplets. Component fermion fields of the multiplet with zero bare mass acquire masses given by $M = \sqrt{2}|a|$ when the $SU(2)$ gauge symmetry is spontaneously broken. These BPS states can be arranged into a vector invariant under $SO(2N_f)$ transformations. Secondly there are fermionic zero modes of the magnetic monopole states in the theory which create BPS states. When the gauge symmetry of the theory is broken, magnetic monopoles appear. For each $SU(2)$ doublet of fermions, these monopoles induce one zero mode. For N_f matter multiplets present, there are $2N_f$ fermion doublets present and thus $2N_f$ fermion zero modes. In the quantum theory these fermionic zero modes form a spinor representation of $SO(2N_f)$. These BPS states therefore transform as a spinor invariant under $SO(2N_f)$ transformations, and the monopoles in the theory can be described using spinors. With spinors in the quantum theory, the symmetry group is then expected to be a universal cover of $SO(2N_f)$. The groups which cover $SO(2N_f)$ universally are the $Spin(2N_f)$ groups.

We now describe the particle spectrum of the quantum theory. When the classical theory is quantized, the magnetic monopoles of the theory acquire electric charges. Via the Witten effect, a 2π rotation of the electric charge operator for these states produces a (gauge) transformation different to identity which has non-trivial topology [168]. For a monopole with magnetic quantum number $n_m = 1$, this transformation has an eigenvalue $e^{i\vartheta}(-1)^H$. The factor $(-1)^H$ can be identitied with the centre of the $SU(2)$ group and also with the chirality operator of the $SO(2N_f)$ spinor. This factor is odd for states in the matter multiplet and even for states in the vector multiplet. Electrically charged states with electric quantum number $n_e \in \mathbb{Z}$ are such that $(-1)^H = 1$ for n_e even and $(-1)^H = -1$ for n_e odd. In addition, when $N_f = 1, 3$, the $SO(2N_f)$ parity ensures that there is a degeneracy between dyons of opposite $SO(2N_f)$ spinor chirality and opposite electric and magnetic charges. This simplifies the particle spectrum in the theories with one or three fundamental matter multiplets. This degeneracy in the spectrum does not occur for $N_f = 2, 4$.

For $N_f \geq 2$, the universal cover of the $SO(2N_f)$ BPS symmetry can be used to label the states in the particle spectrum. Following [171, 190], we label the states in terms of the centres of the $Spin(2N_f)$ symmetry groups.

$N_f = 2$: The symmetry group is $Spin(4) \simeq SU(2) \times SU(2)$. This group has centre $\mathbb{Z}_2 \times \mathbb{Z}_2$. Let representations of the centre be labelled by (η, η'). Then an elementary fermion state is given by $(\eta, \eta') = (1, 1)$. A generically charged fermion state then transforms under the centre $\mathbb{Z}_2 \times \mathbb{Z}_2$ as:

$$(\epsilon, \epsilon') = ((n_e + n_m) \bmod 2, n_e \bmod 2). \tag{5.138}$$

$N_f = 3$: The symmetry group is $Spin(6) \simeq SU(4)$. The centre of this group is $\mathbb{Z}_4$. In this case a generically charged fermion state transforms under the centre $\mathbb{Z}_4$ as:

$$\exp \frac{i\pi}{4}(n_m + 2n_e). \tag{5.139}$$

$N_f = 4$: The symmetry group is $Spin(8)$. This group has centre $\mathbb{Z}_2 \times \mathbb{Z}_2$. In a similar way to the transformations of fermion states in the case when $N_f = 2$, generically charged fermion states transform under the centre $\mathbb{Z}_2 \times \mathbb{Z}_2$ as:

$$(\epsilon, \epsilon') = (n_m \bmod 2, n_e \bmod 2). \tag{5.140}$$

If non-zero $\mathcal{N} = 2$ mass terms are added to the theory, in the form of an $\mathcal{N} = 1$ superpotential as described above, then the $SO(2N_f)$ symmetry is broken. We denote such non-zero masses as m_{N_f}. When $m_{N_f} \neq 0$, the $SO(2N_f)$ symmetry group is broken to $SO(2N_f - 2) \times SO(2)$. The central charge of the theory is then modified and becomes:

$$Z = an_e + a_D n_m + \frac{1}{\sqrt{2}} S_{N_f} m_{N_f}, \tag{5.141}$$

where S_{N_f} is given by Eq. (5.128) with $i = N_f$. This implies that when a assumes the values $a = \pm m_{N_f}/\sqrt{2}$, an elementary fermion state becomes massless. Since the theory to be determined is a low energy Wilsonian effective theory, this behaviour will be important for the determination of the singularity structure of the quantum moduli space.

Electric-Magnetic Duality in $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD

When there are no matter multiplets, $N_f = 0$, in the theory, there is an effective electric-magnetic $SL(2,\mathbb{Z})$ duality present in the theory. The $SL(2,\mathbb{Z})$ duality group acts upon the fields and couplings of the theory. When fundamental matter multiplets are present, the monodromy transformations differ from the $N_f = 0$ case and depend on the masses of the matter multiplets.

The monodromy matrices transform the vector (a, a_D) around singularities of the moduli space. The most simple possibility is when one fermion field has a non-zero bare mass m_{N_f}. Near the point $a \approx a_0$, where $a_0 \equiv \frac{1}{2}\sqrt{2} m_{N_f}$, we expect a singularity to occur in the limit $m_{N_f} \to 0$, in which the fermion state becomes massless. From similar behaviour in $\mathcal{N} = 2$ supersymmetric QED, a logarithmic singularity is expected to result. This is indeed the case, for near the singularity, one has:

$$a \approx a_0 \equiv \frac{m_{N_f}}{\sqrt{2}}, \tag{5.142}$$

$$a_D \approx \frac{i}{2\pi}(a - a_0)\ln(a - a_0) + c, \tag{5.143}$$

where c is an arbitrary constant. The monodromy transformations resulting from a 2π loop about this point on the moduli space is:

$$a \to a' = a, \tag{5.144}$$

$$a_D \to a'_D = a_D + a - \frac{m_{N_f}}{\sqrt{2}}. \tag{5.145}$$

The monodromy transformations in Eqs. (5.144,5.145) are inhomogeneous transformations. This is because the vector (a, a_D) is translated under the monodromy (by the amount $\frac{1}{2}\sqrt{2}m_{N_f}$) as well as being transformed by $SL(2,\mathbb{Z})$ duality transformations. This is unlike action of the monodromies for the $N_f = 0$ case, given in Eqs. (5.71,5.72) of Subsection 5.3.1. The BPS mass bound is consistent with this additional shift since it is also modified by the presence of massive matter multiplets. For $N_f = 0$, such additional translations cannot be part of the monodromy group. To include the translation term in Eq. (5.145) in the monodromy matrix M_s, where the subscript s denotes the singularity it applies to, we describe the action of $M_{m_{N_f}}$ on the column vector $(m_{N_f}/\sqrt{2}, a_D, a)^T$. Then the monodromy matrix $M_{m_{N_f}}$ acts as:

$$\begin{pmatrix} m_{N_f}/\sqrt{2} \\ a_D \\ a \end{pmatrix} \rightarrow \begin{pmatrix} m'_{N_f}/\sqrt{2} \\ a'_D \\ a' \end{pmatrix} = M_{m_{N_f}} \begin{pmatrix} m_{N_f}/\sqrt{2} \\ a_D \\ a \end{pmatrix}, \tag{5.146}$$

where:

$$M_{m_{N_f}} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad M^{-1}_{m_{N_f}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}. \tag{5.147}$$

Furthermore, the action of the monodromy on the central charge Z (and thus the mass M) can be considered. The quantum numbers appearing in Z can be arranged into the row vector $W = (S, n_m, n_e)$. For Z to be invariant under the M_s, the charges contained in W will transform to WM_s^{-1}. The general form of the monodromy matrix M_s will then be of the form:

$$M_s = \begin{pmatrix} 1 & 0 & 0 \\ r & k & l \\ q & n & p \end{pmatrix}, \quad M_s^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ lq - pr & p & -l \\ nr - kq & -n & k \end{pmatrix}, \quad \det(M_s) = 1. \tag{5.148}$$

The monodromy transformations given by M_s permit the mixing of electric and magnetic charges. Under these transformations, the electric and magnetic quantum numbers n_e and n_m can receive translations proportional to n_e and n_m, but not translations proportional to S. This is because S is a global symmetry charge. Conversely, however, S can be translated by amounts proportional to n_e and n_m, which arise from local gauge symmetry. For the specific monodromy transformation $M_{m_{N_f}}$, the row vector of charges W is shifted as follows:

$$M_{m_{N_f}} \; : \; (S, n_m, n_e) \rightarrow (S', n'_m, n'_e) = (S + n_m, n_m, n_e - n_m). \tag{5.149}$$

The formalism of the theory is $SL(2,\mathbb{Z})$ covariant, but the spectrum of particle states is not. There exists further evidence for $SL(2,\mathbb{Z})$ duality in the $N_f = 4$ theory, but we do not describe this further, and refer the reader to the second paper of Seiberg and Witten [171].

Moduli Spaces of $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f \leq 3$

In analogy to the case when $N_f = 0$, the singularities on the moduli space are to be determined before their monodromies are calculated. This assists in elucidating the structure of the moduli space of the theory. The renormalization group behaviour of $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with different values of $N_f \leq 4$ is important in the analysis of Seiberg and Witten of the moduli spaces of these theories. The beta function of these theories in Eq. (5.121) can be integrated (using a standard normalization of g) to give:

$$\frac{1}{\alpha_{N_f}(\mu)} = \frac{(4-N_f)}{2\pi}\ln\left(\frac{\mu}{\Lambda_{N_f}}\right), \tag{5.150}$$

where $\alpha_{N_f} = 4\pi/g^2_{N_f}$ is the inverse squared gauge coupling and Λ_{N_f} is the dynamically generated (energy cutoff) scale at a given value of $N_f \leq 4$, and μ is the energy scale.
If the theory is considered at a scale $\mu < m$ and $N_f - N_f'$ of the fermion states have a mass $m_i = m$ $(i = 1, \ldots, N_f)$, the low energy Wilsonian effective theory will possess only N_f' matter multiplets as effective degrees of freedom. In this case the effective coupling $\alpha_{N_f'}(\mu)$ of theory is given by Eq. (5.150) with N_f and Λ_{N_f} replaced by N_f' and $\Lambda_{N_f'}$. The effective scales of the two theories are related by the requirement that the matching $\alpha_{N_f}(m) = \alpha_{N_f'}(m)$ holds. This matching then implies that the scales of the theories are related via:

$$\Lambda_{N_f'}^{4-N_f'} = m^{N_f-N_f'}\Lambda_{N_f}^{4-N_f}. \tag{5.151}$$

Using the matching given in Eq. (5.151), the scale of theories in the class of $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with $N_f \leq 3$ fundamental matter multiplets can be related. The case of $N_f = 4$ is special as it gives a scale invariant theory which must be treated separately, as will be described below.
Via renormalization group flow, the theory with $N_f = 3$ is equivalent to the theory with

$N_f = 2$ in the limit of one of the matter multiplet bare masses tending to zero. In this way, information about the moduli spaces of theories with $N_f - 1$ multiplets can be obtained from the theory with N_f multiplets. To do this, one can begin with theories in which the bare matter multiplet masses are much greater than the scale Λ. When the masses of the matter multiplets tend to zero, and singularities form, they do so in the large $|u|$ region of the moduli space, which is also the semi-classical region. This makes these singularities readily identifiable. When $|u|$ is small, the theory is effectively the $N_f = 0$ theory, with two singularities corresponding to massless magnetic monopoles and dyons. We now examine the cases individually, and defer the special case of $N_f = 4$ to below.

$N_f = 3$: Let the three matter multiplets have equal masses $m_i = m > \Lambda$. Then the global $Spin(6) \approx SU(4)$ symmetry of the theory is broken to $SU(3) \times U(1)$. Classically there exists a singularity at $a = m/\sqrt{2}$, where all masses vanish, $m_i = 0$. The resulting massless states form a triplet of $SU(3)$. When $|u|$ is small, the moduli space is the same as that of the $N_f = 0$ theory. It has a dynamically generated scale Λ_0 and two singularities at $(n_m, n_e) = (1,0)$ and $(n_m, n_e) = (1,1)$, at which monopoles and dyons become massless. These singular points are $SU(3)$ invariant. As one of the multiplet masses is decreased towards zero, $m \to 0$, the previous global symmetry group of the theory is restored and the massless states form representations of $SU(4)$. This constrains the singularities of the moduli space, since they must combine to allow this. Due to the transformation properties of the massless states under $SU(3)$, these representations can only be a singlet and a quadruplet of $SU(4)$. Given the transformation properties of generic charged particle states under the centre of the symmetry group $Spin(6)$, as described above, these states correspond to massless states whose smallest charges are $(n_m, n_e) = (1,0)$ for the singlet state and $(n_m, n_e) = (2,1)$ for the quadruplet state. Hence there are three singularities on the moduli space for the $N_f = 3$ theory.

$N_f = 2$: When the two matter multiplets present have equal masses $m_i = m > \Lambda$, the global symmetry group $Spin(4) \approx SU(2) \times SU(2)$ is broken to the product $SO(2) \times SO(2)$. Again there is a classical singularity at $a = m/\sqrt{2}$, at which all the masses vanish, $m_i = 0$.

These massless states form representations of one of the $SO(2)$ groups. At small $|u|$, the moduli space again has the same form as the $N_f = 0$ case, and has two singular points at $(n_m, n_e) = (1, 0)$ and at $(n_m, n_e) = (1, 1)$. These singular points transform as singlets under $SO(2) \times SO(2)$. When the limit $m \to 0$ is taken, the previous global symmetry group is restored, and again the massless states of the theory form representations of $Spin(4)$. This constraint on the massless states (that is, the singularities of the moduli space) implies that there are two singularities, which form two different representations of $SU(2) \times SU(2)$. Using the formula for the behaviour of generic charged states under the centre of $Spin(4)$, these singularities occur for the smallest charges satisfying $(n_m, n_e) = (1, 0)$ and at $(n_m, n_e) = (1, 1)$, which transform as spinors in $SO(4)$. Thus there exist four massless states associated with three singularities on the moduli space for the $N_f = 2$ theory.

$N_f = 1$: In this case, the massless theory and the massive theory possess the same $SO(2)$ global symmetry. Arguments analogous to those for the $N_f = 3$ and $N_f = 2$ cases imply that there are three singularities on the moduli space for $m_i = m > \Lambda$. When the limit $m \to 0$ is taken, no additional singularities appear due to the $\mathbb{Z}_3$ symmetry of the u-plane in this case. The transformation properties of a and a_D under the R-symmetry described above indicate that one of the singularities in the zero mass limit is a massless state with charges $(n_m, n_e) = (1, 0)$. The $\mathbb{Z}_3$ symmetry of the moduli space implies that two other singularities exist, with charges $(n_m, n_e) = (1, 1)$ and $(n_m, n_e) = (1, 2)$. When all matter multiplets become massless, the $N_f = 1$ theory still possesses the three singularities above.

Using the same identification as for the $N_f = 0$ theory, in which the moduli space of vacua is considered to be isomorphic to the moduli space for a genus one Riemann surface, the vacuum moduli space can also be specified by a family E_u of elliptic curves [250] in which the holomorphic co-ordinate u appears as a parameter. The variables $a(u)$ and $a_D(u)$ can then be calculated as the periods of these curves along appropriate homology cycles. Knowledge of the singularities and corresponding monodromies on the moduli space then allows the reconstruction of the elliptic curves which describe the moduli space of a Riemann surface with which the moduli space of vacua has been identified.

Firstly, the Seiberg–Witten elliptic curve and the monodromies for the $N_f = 0$ theory has to be described in terms of the new charge normalization conventions, following Seiberg and Witten and introduced at the beginning of this subsection. This will then allow the general characteristics of the $N_f = 0$ curve to act as a guide to the form of the elliptic curves for $N_f \leq 3$. In the new charge normalization convention, the $N_f = 0$ elliptic curve, which is the original Seiberg–Witten elliptic curve, assumes the form:

$$y^2 \;=\; x^3 - ux^2 + \tfrac{1}{4}\Lambda^4 x. \tag{5.152}$$

The monodromy matrices which describe the singularities of the Seiberg–Witten curve, which originally formed the subgroup $\Gamma(2) \subset SL(2,\mathbb{Z})$, now form the subgroup $\Gamma_0(4) \subset SL(2,\mathbb{Z})$ using the new charge normalization, and read:

$$M_\infty \;=\; \begin{pmatrix} -1 & 4 \\ 0 & -1 \end{pmatrix}, \quad M_1 \;=\; \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad M_{-1} \;=\; \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}. \tag{5.153}$$

Following Seiberg and Witten [171], we now analyse the properties and singularities of genus one elliptic curves. Any elliptic curve of genus one can be expressed as a cubic in x:

$$y^2 \;=\; F(x) \;=\; (x - e_1)(x - e_2)(x - e_3), \tag{5.154}$$

which describes the space x which is a double cover of the complex plane with branch points at e_j, $j = 1, 2, 3$, and $x \to \infty$. For the family of elliptic curves E_u, these branch points will in general depend on u. When two branch points of the curve coincide, the curve becomes singular. This type of singularity is referred to as a stable singularity. If more than two branch points of the curve coincide, the curve also becomes singular, but the singularity is not stable. However, a reparameterization of x and y which depends on the holomorphic co-ordinate u can be always be performed in which an unstable singularity becomes stable. The Seiberg–Witten elliptic curve Eq. (5.152) has branch points at $x = 0, \frac{1}{2}(u \pm \sqrt{u^2 - \Lambda^4})$, and at $x \to \infty$. When $u = \pm\Lambda^2$, the two resulting singularities are stable. The singularity at $u \to \infty$ is not stable.

To understand the physical consequences of stable and unstable singularities, let a generic elliptic curve have a stable singularity at $u = 0$. Then the family of elliptic curves with such a singularity, near $u = 0$, can be written as:

$$y^2 \;=\; (x - 1)(x - u^n), \tag{5.155}$$

where $n \in \mathbb{Z}$. The monodromy about u will be conjugate to the matrix T^n, where T, as defined in Eq. (5.54) of Subsection 5.3.1, is a generator of the $SL(2,\mathbb{Z})$ duality group.

A quantity which can be used to classify the singularities of a polynomial in terms of stable and unstable singularities is the discriminant Δ. This is defined as:

$$\Delta = \prod_{i<j}(e_i - e_j)^2, \tag{5.156}$$

where the e_i are the roots of the polynomial. In general, Δ can be written in terms of the coefficients of the polynomial. When two branch points coincide, one obtains $\Delta = 0$, excluding singularities at infinity. For the elliptic curve in Eq. (5.155), near the singularity $u = 0$, $\Delta \sim u^n$. The monodromy at this singularity is conjugate to T^n. At the branch points $u = \pm\Lambda^2$, Δ will be of order unity. The monodromies at these singularities are conjugate to T. In general, the exponent of the monodromy at a stable singularity will be of the same order of the zero of Δ at the singularity.

For unstable singularities, a similar result holds. At large values of u, the branch points of the Seiberg–Witten curve in Eq. (5.152) approximately occur at $x = 0, \Lambda^4/4u, u$ and $x \to \infty$. The singularity at $u \to \infty$ is unstable as more than two branch points coincide in this limit of u. The change of variables $x = x'u$ and $y = y'u^{3/2}$ shifts the branch points at large u to approximately $x = 0, \Lambda/4u^2, 1$ and $x \to \infty$. The singularity in the limit $u \to \infty$ is now stable, and the discriminant in this limit is $\Delta \sim u^{-4}$, which has a monodromy conjugate to T^{-4}. This monodromy corresponds to the monodromy $M_\infty = PT^{-4}$ given in Eq. (5.73) of Subsection 5.3.1, in terms of the original variables (x, y), hence correctly reproducing the behaviour of the Seiberg–Witten elliptic curve at this singular point.

We now turn to the actual construction of the monodromies. Since the charge quantum numbers of the particle spectrum which becomes massless and comprises the singularities of the moduli space is known, the procedure of determining the explicit form of the monodromies is directly analogous to that for the $N_f = 0$ theory.

The monodromy at $u \to \infty$ for the theory with $0 < N_f \leq 4$ matter multiplets can be determined directly from the perturbative beta function given by Eq. (5.121), in the same way as for the $N_f = 0$ theory. The generic result is:

$$M_\infty = PT^{N_f-4}. \tag{5.157}$$

Singularities at finite $|u|$ are in general associated with massless magnetic monopoles of charge $(n_m, n_e) = (1, n_e)$. The monodromies at these points can be obtained by using a dual description of the theory in which magnetic monopoles and the dual gauge field couple in the same way as electric charges couple to the gauge field. Using the one loop beta function of $\mathcal{N} = 2$ supersymmetric QED, the monodromy for k massless magnetic monopoles is given by T^k. To express this result in the dual description of the theory, one can apply the duality transformation $T^{n_e}S$, which transforms a matter multiplet of purely electric charge $(0, 1)$ into a magnetic multiplet of charge $(1, n_e)$, to the monodromy T^k. Hence the monodromy for a point at which k magnetic monopoles of charge $(1, n_e)$ becomes massless is given by $(T^{n_e}S)T^k(T^{n_e}S)^{-1}$. An analogous argument exists which determines the monodromy of a massless state of charge $(2, 1)$ for the $N_f = 3$ theory. The monodromies for the singularities of the moduli space of $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with $N_f \leq 3$ matter multiplets can now be calculated explicitly. In each case, the product of the monodromy matrices at finite $|u|$ gives the monodromy matrix at $u \to \infty$, M_∞. The results can be summarized as follows, in which the values of u where singularities for $N_f < 0$ occur cannot yet be specified, since these values are not known explicitly:

$$N_f = 0 : M_1 = STS^{-1},\ M_{-1} = (T^2S)T(T^2S)^{-1},\ M_\infty = PT^{-4}, \quad (5.158)$$

$$N_f = 1 : STS^{-1},\ (TS)T(TS)^{-1},\ (T^2S)T(T^2S)^{-1},\ M_\infty = PT^{-3}, \quad (5.159)$$

$$N_f = 2 : ST^2S^{-1},\ (TS)T(TS)^{-1},\ M_\infty = PT^{-2}, \quad (5.160)$$

$$N_f = 3 : (ST^2S)T(ST^2S)^{-1},\ ST^4S^{-1}, M_\infty = PT^{-1}, \quad (5.161)$$

where the form of M_∞ for each value of N_f satisfies Eq. (5.157).

The Elliptic Curve for $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f = 0$

The $N_f = 0$ elliptic curve is the Seiberg–Witten elliptic curve given in Eq. (5.152) in the new charge normalization conventions. There is no distinction between the massless and massive cases of the theory as there are no matter multiplets present. Some properties of the Seiberg–Witten curve are also present for the curves with non-zero N_f. These include the following:

i. The elliptic curves for $N_f > 0$ are anticipated to be of the form $y^2 = F(x, u, \Lambda)$, where F is a polynomial of degree no higher than three in x and u. This property will ensure the correct monodromy matrices for the singularities in the theories with $N_f > 0$.

ii. The term which is cubic in x in the polynomial F has the form $F_0 = x^2(x - u)$. The singularity in the limit $|u| \to \infty$ must be present for all values of N_f, and the monodromy at infinity will be that derived from the one loop beta function, namely $M_\infty = PT^{N_f - 4}$. For this to occur, one of the branch points of the polynomial F must tend to infinity as $|u| \to \infty$. If this singularity is to remain stable, then the other two branch points must coincide. This constrains the form of the curve, and through a reparameterization of x, the term F_0 can be put into the form $F_0 = x^2(x - u)$. This can also be expressed as the requirement that in the limit $\tau \to i\infty$, which is the weak coupling limit, one should recover the curve $y^2 = F_0$.

iii. One can assign $U(1)_r$ charges to the variables x and u such that $y^2 = F$ is invariant under $U(1)_r$ transformations. If x and u are each assigned a $U(1)_r$ charge of 2, and the scale Λ has $U(1)_r$ charge 2, property (i) dictates that F has a $U(1)_r$ charge of 12. If y is further assigned a $U(1)_r$ charge of 6, then the curve $y^2 = F$ is $U(1)_r$ invariant.

iv. The polynomial F can always be written in the form $F = F_0 + \frac{1}{4}\Lambda^4 x$, where F_0 is defined in property (ii). This is due to the $U(1)_r$ charge assignments in the $N_f = 0$ theory.

The Elliptic Curve for $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f = 1$

We first describe the $N_f = 1$ theory with massless multiplets. In this case, the instanton amplitude given by Eq. (5.135) is proportional to Λ_1^3, where $\Lambda_1 = \Lambda_{N_f}$ is the dynamically generated scale of the $N_f = 1$ theory. When $N_f \geq 1$, the instanton amplitude factor is odd under the ρ parity transformation defined in Eq. (5.129). Since ρ parity is a symmetry of the theory, only odd powers of the factor Λ_1^3 should appear in the $N_f = 1$

elliptic curve. If Λ_1 is assigned the $U(1)_r$ charge according to property (iii) of the $N_f = 0$ curve, and F has a $U(1)_r$ charge of 12, then only a factor of Λ_1^6 may appear in F. Hence the unique elliptic curve for the massless theory which obeys the properties listed for the Seiberg–Witten elliptic curve is:

$$y^2 \;=\; x^2(x-u) + t\Lambda_1^6, \tag{5.162}$$

where t is a constant which may be absorbed by a rescaling of Λ_1 as $\tilde{\Lambda}_1^6 = t\Lambda_1^6$. The discriminant of the set of massless $N_f = 1$ curves Eq. (5.162) is given by:

$$\Delta \;=\; \tilde{\Lambda}_1^6(4u^3 - 27\tilde{\Lambda}_1^6), \tag{5.163}$$

which has three zeros. The zeros of Δ are related to one another by the $\mathbb{Z}_3$ symmetry of the u-plane. The monodromies associated with these singularities are matrices conjugate to T. These matrices, and the monodromy matrix for the singularity at large $|u|$ derived from the elliptic curve are consistent with the monodromies on the moduli space given in Eq. (5.159).

For the theory with $N_f = 1$ massive multiplets, the curve of the massless theory given in Eq. (5.162) with the aforementioned rescaling of Λ_1 can be generalized to give:

$$y^2 \;=\; x^2(x-u) + \tfrac{1}{4}m\Lambda_1^3 x - \tfrac{1}{64}\Lambda_1^6, \tag{5.164}$$

where m is the mass of the fundamental matter multiplet.

The Elliptic Curve for $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f = 2$

For the massless $N_f = 2$ theory, the instanton factor given in Eq. (5.135) is proportional to Λ_2^2. When the two matter multiplets present are massless, only even powers of the factor Λ_2^2 may appear in the polynomial F for the $N_f = 2$ elliptic curve. As in the $N_f = 1$ case, the dynamically generated scale Λ_2 has a $U(1)_r$ charge of 2. The properties of the Seiberg–Witten elliptic curve then constrain the curve of the massless $N_f = 2$ theory to be:

$$y^2 \;=\; x^2(x-u) + (ax + bu)\Lambda_2^2, \tag{5.165}$$

where a and b are constants. The singularity structure of the moduli space for this theory is such that two singularities arise from magnetic monopoles and become massless at finite

$|u|$. From Eq. (5.160), the monodromy matrices at these singular points are conjugate to T^2. Hence the discriminant of the elliptic curve Eq. (5.162) should have second order zeros at these singularities. This restricts the form of the discriminant Δ, from which the constants a and b can be determined. In analogy with the $N_f = 1$ case, these constants can be absorbed into Λ_2 via a rescaling, which gives the scale $\tilde{\Lambda}_2$. After this has been done, the family of elliptic curves which describe the moduli space of the massless $N_f = 2$ theory can be written as:

$$y^2 = (x^2 - \tilde{\Lambda}_2^4)(x - u). \tag{5.166}$$

The elliptic curve exhibits $\mathbb{Z}_2$ symmetry in the u-plane, which is consistent with the $\mathbb{Z}_2$ symmetry which relates the finite $|u|$ singularities on the moduli space of this theory described previously.

When the two matter multiplets of the theory are massive, the elliptic curve in Eq. (5.166) can be generalized to the massive case and then assumes the form:

$$y^2 = (x - \frac{1}{64}\Lambda_2^4)(x - u) + m_1 m_2 \tilde{\Lambda}_2^2 x - \frac{1}{64}(m_1^2 + m_2^2)\tilde{\Lambda}_2^4, \tag{5.167}$$

where m_1 and m_2 are the bare masses of the two fundamental matter multiplets in the theory.

The Elliptic Curve for $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f = 3$

When there are $N_f = 3$ massless matter multiplets coupled to $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD, there are two singularities on the u-plane at finite $|u|$. As expressed in Eq. (5.161), these singularities have monodromies which are conjugate to T^4 and T, respectively. These monodromy matrices are not related via any symmetry which acts on the u-plane. If it is supposed that the singularity with monodromy matrix conjugate to T^4 is at $u = 0$, then the discriminant should have an order four zero at the singular point $u = 0$. The assignments of $U(1)_r$ and ρ parity charges in the theory, involving the instanton factor Eq. (5.135) for $N_f = 3$, then serve to restrict the form of the possible elliptic curves. Together with the properties of the Seiberg–Witten elliptic curve, these constraints imply that the massless $N_f = 3$ theory has a moduli space described by an elliptic curve given by:

$$y^2 = a\Lambda_3^2 + bu^2x + cux^2 + x^3, \tag{5.168}$$

where a, b and c are constants. The constant b must be non-zero, $b \neq 0$, otherwise the elliptic curve is singular for all values of u. As in the previous cases, these constants can be absorbed into the definition of Λ_3 via a rescaling, resulting in the scale $\tilde{\Lambda}_3$. Further rescaling of y in Eq. (5.168) and property (ii) of the $N_f = 0$ curve permits the elliptic curve for the massless N_f theory to be expressed as:

$$y^2 = x^2(x-u) + \tilde{\Lambda}_3^2(x-u)^2. \tag{5.169}$$

The massless $N_f = 3$ elliptic curve Eq. (5.169) can be generalized to the case of massive matter multiplets. The result for the elliptic curve which describes the moduli space of the massive $N_f = 3$ theory is then:

$$\begin{aligned} y^2 &= x^2(x-u) - \frac{1}{64}\Lambda_3^2(x-u)^2 - \frac{1}{64}(m_1^2+m_2^2+m_3^2)\Lambda_3^2(x-u) + \\ &\frac{1}{4}m_1m_2m_3\Lambda_3 x - \frac{1}{64}(m_1^2m_2^2+m_2^2m_3^2+m_1^2m_3^1)\Lambda_3^2, \end{aligned} \tag{5.170}$$

where m_1, m_2 and m_3 are the bare masses of the fundamental matter multiplets.

Determination of the Low Energy Effective Action for $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f \leq 3$

We have described how the elliptic curves whose moduli spaces specify the physical moduli space of vacua for the $\mathcal{N} = 2$ SQCD theories with $N_f \leq 3$ massive and massless fundamental matter multiplets have been deduced in Seiberg–Witten theory [171]. The methods employed to extract these elliptic curves have used symmetries of the theory and properties of the moduli spaces involved.

A procedure directly analogous to that for Seiberg–Witten theory (that is, the elliptic curve for the massless $N_f = 0$ theory) [171] can be implemented to explicitly determine the variables a and a_D for each case when $0 < N_f \leq 3$ and $N_f \in \mathbb{Z}$. Through the identification of the moduli spaces of vacua with the corresponding moduli spaces of specific genus one Riemann surfaces, described by the aforementioned elliptic curves, the variables a and a_D are proposed to be periods of the elliptic curves, given by the contour integrals:

$$a = \oint_{\gamma_1} \lambda, \quad a_D = \oint_{\gamma_2} \lambda, \tag{5.171}$$

where γ_1 and γ_2 are appropriate homology one-cycles with unit intersection number. The integrand λ is a holomorphic differential one-form which satisfies the differential equation:

$$\frac{d\lambda}{du} = \frac{\sqrt{2}}{8\pi}\frac{dx}{y}. \tag{5.172}$$

The metric on the moduli space in each case is then given by $\mathrm{Im}(\tau)$, where τ is the period matrix of the Riemann surface described by the elliptic curve $y^2 = F(x, u, \Lambda)$. The period matrix τ is identified with the complexified coupling constant $\tau(u)$ via the important hypothesis that the moduli space of vacua is isomorphic to the moduli space of the Riemann surface, which makes possible the analysis leading to the determination of the low energy effective prepotential $\mathcal{F}$. In terms of the variables a and a_D, the matrix τ is given by:

$$\tau = \frac{da_D/du}{da/du}. \tag{5.173}$$

For the case of gauge group $SU(2)$ and a four dimensional gauge theory, the matrix τ is given by a complex scalar function. The prepotential $\mathcal{F}$ and the low energy Wilsonian effective action of $\mathcal{N} = 2$ $SU(2)$ SQCD follow using the methods already described in Subsection 5.3.1.

Moduli Space of $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f = 4$

The case when there are $N_f = 4$ fundamental matter multiplets in the theory has a vanishing one loop beta function. This can be seen from the beta function Eq. (5.5) of Section 5.2 in Chapter 4 with $N_f = 4$ substituted. The resulting classical theory is scale invariant. The exact $N_f = 4$ quantum perturbative beta function is zero to all orders in g. Seiberg and Witten propose an argument involving the consequences of the appearance of the instanton factor Eq. (5.135) in the beta function and conclude that the quantum non-perturbative $N_f = 4$ beta function is also zero to all orders in g. If this were not the case, the metric on the moduli space would also not be positive definite. The quantum theory is exactly scale invariant. This scale invariance also implies that the classical and quantum moduli spaces are identical.

For this theory, the complexified coupling constant τ is a dimensionless coupling, due to the absence of dimensional transmutation. In physical terms this means that τ is not a running coupling, and has the same value at all energy scales. There are no quantum perturbative or non-perturbative corrections to τ and so the variables a and a_D also receive no quantum corrections. Therefore, according to this reasoning, these variables maintain their classical form in the quantum theory, namely:

$$a = \frac{1}{2}\sqrt{2u}, \quad a_D = \tau a. \tag{5.174}$$

This implies that the co-ordinate used to parameterize the quantum moduli space, namely the quantum modulus $u = u_{qu}$, is taken to be equal to the co-ordinate on the classical moduli space in this case, $u = u_{cl} = u_{qu}$. We shall return to the identification in Eq. (5.174) of the quantum variables with their classical counterparts in Section 6.5 of Chapter 6. When all four matter multiplets in the theory are massless, there is only one monodromy on the moduli space. This is associated with the origin of the u-plane. When some or all of the matter multiplets are massive, the theory exhibits different behaviour. If some of the masses m_i, $i = 1, \ldots, 4$ are taken to be infinitely large, the quantum $N_f = 4$ moduli space should reproduce the quantum moduli spaces of the asymptotically free theories with $N_f \leq 3$ matter multiplets. If there are masses m_i, $i = n+1, \ldots, 4$, as m_i tends to infinity, the correct scaling limit to be taken to reach these other theories is $\tau \to i\infty$, with u and the quantity Λ_n^{4-n} held fixed. The 'scale' Λ_n^{4-n} is given by:

$$\Lambda_n^{4-n} \sim \sqrt{q}\prod_i m_i, \quad q \equiv e^{i\pi\tau}, \tag{5.175}$$

where τ is the dimensionless complexified coupling constant in the conventions of [171]:

$$\tau = \frac{\vartheta}{\pi} + \frac{8\pi i}{g^2}, \tag{5.176}$$

where the new charge renormalization conventions have been used. The low energy effective theory then has $N_f = n$ matter multiplets and a dynamically generated scale Λ_n. Hence, by treating the scale invariant $N_f = 4$ theory as having more than four matter multiplets, an effective definition of an energy scale in the theory can be found.

Seiberg and Witten consider the cases where differing numbers of matter multiplets are massless in the $N_f = 4$ theory. They also consider those cases where massive matter multiplets have degenerate masses. Each singular point of the u-plane can be weighted by the number of massive matter multiplets becoming massless at that point. The elliptic curve $y^2 = F(x, u)$ which describes the low energy effective theory is such that the discriminant of $F(x, u)$ (with respect to x) is of degree six in u. The singularities of the u-plane are the zeros of this discriminant and so the total weighted number of singularities on the u-plane is six. Thus if all matter multiplets of the theory have unequal masses, there are six singularities each of weight one. When there are matter multiplets with degenerate masses, the singularities associated with them can be combined into one singularity of higher weight. The number of singularities on the $N_f = 4$ moduli space and the global symmetry of the theory depends critically on the number of massless matter multiplets present. Below we summarize the results of the analysis of the various cases of massless, degenerately massive and massive $N_f = 4$ matter multiplets, as given in [171]:

i. $m_i = (m, 0, 0, 0)$: The global symmetry is $SU(4) \times U(1)$. Two singularities of weight one and one singularity of weight four, for which the four massless particles transform in the fundamental representation of $SU(4)$.

ii. $m_i = (m, m, m, m)$: The global symmetry is also $SU(4) \times U(1)$. Two singularities of weight one and one singularity of weight four, for which the four massless particles also transform in the fundamental representation of $SU(4)$.

iii. $m_i = (m, m, 0, 0)$: The global symmetry is $SU(2) \times SU(2) \times SU(2) \times U(1)$. Three singularities each of weight two. The massless particles transform in a doublet of one of the $SU(2)$ factors.

iv. $m_i = (m_1, m_2, 0, 0)$: The global symmetry is $SU(2) \times SU(2) \times U(1) \times U(1)$. Two singularities of weight oneand two singularities of weight two. Massless particles associated with weight one singularities exist in fundamental $SU(4)$ representations, and those for the weight two singularities exist in doublets of one of the $SU(2)$ factors.

v. $m_i = (m_1, m_1, m_2, m_2)$: The global symmetry is also $SU(2) \times SU(2) \times U(1) \times U(1)$. Again two singularities of weight one and two of weight two. The massless particles for the weight two singularities again exist in doublets of one of the $SU(2)$ factors.

vi. $m_i = (m, m, m, 0)$: The global symmetry is $SU(3) \times U(1) \times U(1)$. Three singularities of weight one and one singularity of weight three. The massless particles associated with the weight three singularity exist in fundamental $SU(3)$ representations.

Seiberg and Witten argue that the cases (iv) and (v), and also (i) and (ii), point to evidence of triality in the theory. The moduli spaces for the theories with multiplet masses (iv) and (v) are proposed to be related via this triality. In addition to electric-magnetic duality, there exists a map which transforms τ to itself. Strictly, this is the $\mathbb{S}_3$ automorphism of the full $Spin(8)$ global symmetry group, which acts upon the four masses of the N_f matter multiplets. We note further that the theory with N_f matter multiplets with zero bare masses is conformally invariant.

The singularity structure for the $N_f = 4$ moduli space is dependent on the masses of the matter multiplets present. Similarities between some of the cases (i)–(vi) above and the moduli spaces of $\mathcal{N} = 2$ $SU(2)$ SQCD with $N_f \leq 3$ will be important for determining the elliptic curve description of the $N_f = 4$ moduli space.

Moduli Space of Mass Deformed $\mathcal{N} = 4$ Supersymmetric $SU(2)$ Yang–Mills Gauge Theory

To assist in the determination of the low energy effective action for $\mathcal{N} = 2$ supersymmetric QCD with $N_f = 4$ fundamental matter multiplets, Seiberg and Witten also consider $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory in [171]. These theories are related since $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory coupled to a massless matter multiplet in the *adjoint* representation gives $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory. When the adjoint matter multiplet has a non-zero bare mass, this theory is referred to as mass deformed $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory. The mass deformed $\mathcal{N} = 4$ theory is equivalent to $\mathcal{N} = 2$ supersymmetric QCD with $N_f = 4$ fundamental matter multiplets. Furthermore, both $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory and $\mathcal{N} = 2$ supersymmetric QCD with $N_f = 4$ fundamental matter multiplets have dimensionless coupling constants and are candidates for S-dual or $SL(2,\mathbb{Z})$ invariant theories. Hence it is relevant to describe mass deformed $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory.

The moduli space of the $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory is very similar to that of $\mathcal{N} = 2$ supersymmetric QCD with massless $N_f = 4$ fundamental matter multiplets. There is a dimensionless coupling constant, $\tau_{\mathcal{N}=4}$, given by:

$$\tau_{\mathcal{N}=4} = \frac{\vartheta}{2\pi} + \frac{4\pi i}{g^2}, \tag{5.177}$$

which is identical to the classical complexified gauge coupling given in Eq. (2.17) in Section 2.2 of Chapter 2. Since the theory is scale invariant, the classical and quantum $\mathcal{N} = 4$ moduli spaces are assumed to be identical. The metric on the moduli space, given by $\mathrm{Im}(\tau_{\mathcal{N}=4})da d\bar{a}$, is taken to receive no quantum perturbative or non-perturbative corrections. As described in Section 4.3 of Chapter 4, the vanishing $\mathcal{N} = 4$ beta function is also perturbatively and non-perturbatively exact. The classical and the quantum versions of the theory are taken to have vacuum moduli spaces which can be described in terms of the classical variables a and a_D, given by:

$$a = \sqrt{2u}, \quad a_D = \tau a. \tag{5.178}$$

As shall be described in Section 6.5 of Chapter 6, the assumptions leading to this identification produces some discrepancies when compared with conventional field theoretic calculations.

No assumption is made in the analysis of Seiberg and Witten regarding S-duality of the $\mathcal{N} = 4$ theory. Instead, they put forward evidence of $SL(2,\mathbb{Z})$ invariance in the theory. Below we briefly outline the somewhat lengthy methods which were used to determine the elliptic curves for the mass deformed $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory and the $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with $N_f = 4$ fundamental matter multiplets.

Elliptic Curves for the Massless $\mathcal{N} = 4$ Supersymmetric and $\mathcal{N} = 2$ Supersymmetric Scale Invariant Theories

Following Seiberg and Witten, we now describe some generalities regarding $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory deformed by a massless fundamental matter multiplet, and $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with $N_f = 4$ massless fundamental matter multiplets. Given the classical form of the variables a and a_D for these theories in Eqs. (5.174,5.178), these formulae also apply in the quantized theories, and are exact. Seiberg and Witten make use of the properties of the Weierstrass function, denoted $\wp$, to deduce the form of the elliptic curves which will specify the particular differential forms required. These differential forms are chosen such that they have periods given by $(\partial a_D/\partial u, \partial a/\partial u)$, with a and a_D given by the exact formulae Eqs. (5.174,5.178), and are given by:

$$\omega_{\mathcal{N}=4} = \frac{\sqrt{2}}{8\pi}\frac{dx}{y}, \tag{5.179}$$

$$\omega_{N_f=4} = \frac{\sqrt{2}}{4\pi}\frac{dx}{y}, \tag{5.180}$$

where $\omega_{\mathcal{N}=4}$ and $\omega_{N_f=4}$ are associated with the $\mathcal{N} = 4$ theory and the $\mathcal{N} = 2$ theory with $N_f = 4$ flavours, respectively.

A genus one curve E and an associated differential form with periods given by multiples of $(\tau, 1)$, where τ is the periods matrix for E, can be readily specified. If E is taken to be the quotient of the complex plane $z \in \mathbb{C}$ by the lattice generated by π and $\pi\tau$, then the differential form $\omega_0 = dz$ will have periods π and $\pi\tau$. The algebraic form of the curve E can be found using the aforementioned Weierstrass function $\wp$. This special function has the following properties, as a function of the z-plane:

$$\wp(z) = \wp(z+1) = \wp(z+\tau) = \wp(-z). \tag{5.181}$$

The function $\wp$ has one singularity on E, which is a double pole at the origin. Now let $x_0 = \wp(z)$ and $y_0 = \wp'(z)$. Then one has:

$$y_0^2 = 4x_0^3 - g_2(\tau)x_0 - g_3(\tau), \tag{5.182}$$

where the functions g_2 and g_3 are related to Eisenstein series G_4 and G_6, which define modular forms of weight four and six, respectively:

$$g_2 = \frac{60}{\pi^4}G_4(\tau) = \frac{60}{\pi^4}\sum_{m,n\in\mathbb{Z}}\frac{1}{(m\tau+n)^4}, \tag{5.183}$$

$$g_4 = \frac{140}{\pi^6}G_6(\tau) = \frac{140}{\pi^6}\sum_{m,n\in\mathbb{Z}}\frac{1}{(m\tau+n)^6}, \tag{5.184}$$

where the integers m,n are non-zero. Given that x_0 and y_0 are related via the equation $y_0 = dx_0/dz$, then $\omega_0 = dz$ can be expressed as $\omega_0 = dx_0/y_0$. If, further, one changes variables to $x = x_0 u$ and $y = \frac{1}{2} y_0 u^{3/2}$, then the two differential forms $\omega_{\mathcal{N}=4}$ and $\omega_{N_f=4}$ in Eqs. (5.180,5.179) can be written as:

$$\omega_{\mathcal{N}=4} = \frac{\sqrt{2/u}}{4\pi}\omega_0 = \frac{\sqrt{2}}{8\pi}\frac{dx}{y}, \tag{5.185}$$

$$\omega_{N_f=4} = \frac{\sqrt{2/u}}{2\pi}\omega_0 = \frac{\sqrt{2}}{4\pi}\frac{dx}{y}, \tag{5.186}$$

and the elliptic curve Eq. (5.182) becomes:

$$y^2 = x^3 - \frac{1}{4}g_2(\tau)xu^2 - \frac{1}{4}g_3(\tau)u^3. \tag{5.187}$$

The form of this change of variables was chosen for consistency with the properties of the Seiberg–Witten elliptic curve which are postulated to extend to the other theories with $0 < N_f \leq 4$ matter multiplets and the required asymptotic behaviour of the curve (in particular, in the weak coupling limit $\tau \to i\infty$). In this construction of the massless elliptic curves, the differential forms $\omega_{\mathcal{N}=4}$ and $\omega_{N_f=4}$ now have periods given by $\sqrt{u/8}(1,\tau)$ and $\sqrt{u/2}(1,\tau)$, and the correct exact formulae for the variables a and a_D in Eqs. (5.174,5.178) are recovered.

We note that the coefficients of the elliptic curves Eqs. (5.182,5.187) are modular forms. This is because the metric on the classical moduli space for these theories is invariant under the modular group $SL(2,\mathbb{Z})$, and is thus S-dual. This fact can be regarded as evidence for S-duality in these theories.

The elliptic curve Eq. (5.187) may also be factorized in terms of Θ functions expressed as functions of the exponentiated complexified (dimensionless) coupling constant $q = e^{2\pi i\tau}$. The Θ functions are modular forms of weight $\frac{1}{2}$, and relate the three roots of the cubic polynomial in Eq. (5.187) when $y^2 = 0$ is set; we denote the three resulting roots as e_i, $i = 1,2,3$. The elliptic curve E given by Eq. (5.187) possesses spin structures which are given by the roots e_i. For the $\mathcal{N} = 4$ theory, the spin structures are not significant. For the $\mathcal{N} = 2$ theory with $N_f = 4$ matter multiplets, the S-duality group $SL(2,\mathbb{Z})$ acts upon the spin structure and the conjectured $Spin(8)$ triality group of the theory. The modular group acts upon both of these objects via permutation, and therefore $Spin(8)$

triality permutes the roots e_i of the massless elliptic curve Eq. (5.187).

The Elliptic Curve for $\mathcal{N}=4$ Supersymmetric $SU(2)$ Yang–Mills Gauge Theory

The $\mathcal{N}=4$ supersymmetric $SU(2)$ Yang–Mills gauge theory can be regarded as $\mathcal{N}=2$ supersymmetric $SU(2)$ QCD coupled to one massless adjoint matter multiplet. If the adjoint multiplet acquires a non-zero bare mass m, that is, if the $\mathcal{N}=4$ theory is mass deformed, then the $\mathcal{N}=4$ supersymmetry is explicitly broken to $\mathcal{N}=2$ supersymmetry. We now describe the singularity structure of the moduli space of the mass deformed $\mathcal{N}=4$ theory, as proposed by Seiberg and Witten. This will assist in the determination of the elliptic curve proposed to describe the moduli space. There exists one singularity at weak coupling, for which $|q| \ll 1$, and $m \neq 0$. This singularity arises from a component field of the massive adjoint multiplet becoming massless, and at this point $u = \frac{1}{4}m^2$. Since the presence of a non-zero bare mass breaks the $\mathcal{N}=4$ supersymmetry, the theory develops into a strongly coupled $\mathcal{N}=2$ supersymmetric $SU(2)$ Yang–Mills gauge theory at energy scales of order $\Lambda_0 \approx q^{1/4}m$. This theory is amenable to the analysis of Seiberg and Witten in [170] and is known to have two singularities at finite u arising from massless monopoles and dyons. All of the three singularities so far uncovered have monodromy matrices conjugate to T^2. At strong coupling, and in the absence of S-duality (which Seiberg and Witten do not assume), the three singularities each arise from a multiplet becoming massless.

Using this fact and the conditions on the elliptic curve in the massless limit $m \to 0$, a first approximate form of the curve can be deduced. Applying the restrictions on the curve in the weak coupling limit, which should be smooth, this first approximation can be refined. This further constrains the mass dependence of the coefficients in the curve. Finally, Seiberg and Witten consider the residues of the differential form λ at the singularities of the elliptic curve. This completely fixes the mass dependence of coefficients in the curve, and leads to the following elliptic curve proposed to describe the moduli space of $\mathcal{N}=4$ supersymmetric $SU(2)$ Yang–Mills gauge theory mass deformed by an adjoint matter multiplet of bare mass m:

$$y^2 \;=\; (x - e_1\tilde{u} - \frac{1}{4}e_1^2m^2)(x - e_2\tilde{u} - \frac{1}{4}e_2^2m^2)(x - e_3\tilde{u} - \frac{1}{4}e_3^2m^2),, \tag{5.188}$$

where $\tilde{u}$ is given by:

$$\tilde{u} = u - \frac{1}{8}e_1 m^2. \tag{5.189}$$

In the elliptic curve Eq. (5.188), the quantities e_i, $i = 1, 2, 3$ are the roots of the cubic polynomial obtained by setting $y^2 = 0$ in Eq. (5.187), which describes the massless $\mathcal{N} = 4$ theory, as previously stated. The elliptic curve Eq. (5.188) for the mass deformed $\mathcal{N} = 4$ theory has modular forms as coefficients, and so it is also invariant under the modular group $SL(2, \mathbb{Z})$. Since this elliptic curve is proposed to describe the quantum moduli space of the theory, this modular invariance or S-duality is valid in the strong coupling regime. This is a hitherto unknown property of the $\mathcal{N} = 4$ theory, which can be regarded as a new test of the conjectured exact extended electric-magnetic duality (S-duality) of this theory. However, this $SL(2, \mathbb{Z})$ invariance does not appear in the weak coupling limit $\tau \to i\infty$ of the theory. In this limit, u is the appropriate co-ordinate of the moduli space, which is not a modular form, whereas the renormalized co-ordinate $\tilde{u}$ is, and is thus $SL(2, \mathbb{Z})$ invariant. Furthermore, in the limit $\tau \to p/q$, $p, q \in \mathbb{Z}$ and $m \to \infty$, the $\mathcal{N} = 4$ theory is broken to an $\mathcal{N} = 2$ Yang–Mills theory, which does not possess $SL(2, \mathbb{Z})$ invariance. In this case, the $SL(2, \mathbb{Z})$ modular group permutes the possible scaling limits which lead to the $\mathcal{N} = 2$ theory.

The Elliptic Curve for $\mathcal{N} = 2$ Supersymmetric $SU(2)$ QCD with $N_f = 4$

Using and extending the techniques previously applied to $\mathcal{N} = 2$ theories with $0 < N_f \leq 3$ matter multiplets, Seiberg and Witten proposed an elliptic curve which describes the quantum moduli space of mass deformed $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory. As $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD with $N_f = 4$ matter multiplets is a similar theory when there are non-zero bare masses present, the determination of the elliptic curve for this theory proceeds in an analogous way.

Seiberg and Witten initially consider the special case when the four matter multiplets have masses $m_i = (m, m, 0, 0)$, which includes two degenerate masses. This scenario was described as case (iii) of the moduli space of the $N_f = 4$ theory above. There are three singularities on the moduli space when $m_i = (m, m, 0, 0)$. This implies that each of the singularities has a monodromy matrix conjugate to T^2. One can recognize this situation as being precisely that of the $\mathcal{N} = 4$ theory described above, and this therefore immediately leads to a first approximation for the required elliptic curve. For the case when the masses are $m_i = (m, m, 0, 0)$, using information from the weak coupling limit, the mass dependence of the special $N_f = 4$ elliptic curve readily follows.

The general $N_f = 4$ elliptic curve will be one in which the masses of the four matter multiplets are arbitrary. By imposing the global $SO(8)$ symmetry of the theory, using the limit to the special case $m_i = (m, m, 0, 0)$, smoothness of the limit $m_i \to \infty$ for any i, and $U(1)_r$ charge assignations, the form of the general $N_f = 4$ elliptic curve can be obtained. The constant coefficients of the curve, which are functions of the cubic roots e_i, are further fixed by considering the other special cases for the matter multiplet masses and the resulting zeros of the discriminant of the curve. Specifically, the cases when $m_i = (m_1, m_2, 0, 0)$, $m_i = (m_1, m_1, m_2, m_2)$ and $m_i = (m_1, m_1, m_3, -m_3)$, are considered. These indirect arguments lead to the final result for the $N_f = 4$ elliptic curve with arbitrary matter multiplet masses, which has the form:

$$y^2 = W_1W_2W_3 + A[W_1T_1(e_2 - e_3) + W_2T_2(e_3 - e_1) + W_3T_3(e_1 - e_2)] - A^2N, \quad (5.190)$$

where $(i = 1, 2, 3)$:

$$W_i = x - e_i\tilde{u} - e_i^2 R, \quad (5.191)$$

$$\tilde{u} = u - \frac{1}{2}e_1 R, \quad (5.192)$$

$$R = \frac{1}{2}\sum_i m_i^2, \quad (5.193)$$

$$T_1 = \frac{1}{12}\sum_{i>j} m_i^2 m_j^2 - \frac{1}{24}\sum_i m_i^4, \quad (5.194)$$

$$T_2 = -\frac{1}{2}\prod_i m_i - \frac{1}{24}\sum_{i>j} m_i^2 m_j^2 + \frac{1}{48}\sum_i m_i^4, \quad (5.195)$$

$$T_3 = \frac{1}{2}\prod_i m_i - \frac{1}{24}\sum_{i>j} m_i^2 m_j^2 + \frac{1}{48}\sum_i m_i^4, \quad (5.196)$$

$$N = \frac{3}{16}\sum_{i>j>k} m_i^2 m_j^2 m_k^2 - \frac{1}{96}\sum_{i\neq j} m_i^2 m_j^4 + \frac{1}{96}\sum_i m_i^6. \quad (5.197)$$

The elliptic curve Eq. (5.190) is fully $SL(2, \mathbb{Z})$ modular invariant, a symmetry which permutes the roots e_i, if it is combined with the $Spin(8)$ triality symmetry, which permutes the $SO(8)$ invariants T_i. Since both of these symmetries are only conjectured to exist in the theory, the form of the massive $N_f = 4$ elliptic curve does not provide proof of S-duality in the theory. This theory cannot be exactly $SL(2,\mathbb{Z})$ invariant since the modular invariance of the elliptic curve is only an effective one, which exists only when $Spin(8)$ triality is also present. The form of the elliptic curve does however support the presence of these symmetries.

A consistency check for the elliptic curve Eq. (5.190) is provided by renormalization group flow from $N_f = 4$ to $N_f < 4$. The elliptic curves for the $N_f \leq 3$ theories are checked by Seiberg and Witten in [171] by taking the weak coupling limit $\tau \to i\infty$ in combination with the mass limit $m_{N_f} \to \infty$ with all other masses fixed. In this limit, the $N_f = 4$ elliptic curve reproduces the elliptic curve for the case of $N_f = 3$ arbitrarily massive matter multiplets. Then further renormalization group flow permits one to recover the original Seiberg–Witten elliptic curve for the $N_f = 0$ theory using this limit.

In [171], Seiberg and Witten also derive the mass deformed $\mathcal{N} = 4$ and $N_f = 4$ SQCD elliptic curves using a condition on the differential form λ. This is the residue condition, which originates from the fact that the variables a and a_D are translated by amounts dependent on the bare masses m_i under the monodromy transformations about the singularities on the moduli spaces of these theories. These translations are given by integral linear combinations of $m_i/\sqrt{2}$. These constants impose the following condition on the residues of λ:

$$\mathrm{Res}\,\lambda \;=\; \sum_i \frac{n_i m_i}{2\pi i\sqrt{2}}, \quad n_i \in \mathbb{Z}. \tag{5.198}$$

By utilizing this condition and lengthy complex and cohomological analysis, the correct elliptic curves Eqs. (5.188,5.190) for the scale invariant theories can be derived. This verifies that the residues of the elliptic curves do indeed obey the residue condition Eq. (5.198), thus providing an internal consistency check of the elliptic curves.

Given the elliptic curves for the mass deformed $\mathcal{N} = 4$ theory and $\mathcal{N} = 2$ SQCD with $N_f = 4$ matter multiplets, the procedure for calculating the variables $a(u)$ and $a_D(u)$ is the same as that used for the theories with $N_f \leq 3$ matter multiplets. The periods of the differential forms $\omega_{\mathcal{N}=4}$ and $\omega_{N_f=4}$ about appropriate homology one-cycles will enable one to calculate the period matrices for these theories, which is again identified with the complexified coupling constant of these theories:

$$\frac{da(u)}{du} = \oint_{\gamma_1} \omega, \quad \frac{da_D(u)}{du} = \oint_{\gamma_2} \omega, \tag{5.199}$$

where ω represents either $\omega_{\mathcal{N}=4}$ or $\omega_{N_f=4}$, as defined in Eqs. (5.179,5.180), respectively, depending on the theory under consideration, and τ is given by:

$$\tau = \frac{da_D(u)/du}{da(u)/du}. \tag{5.200}$$

The metrics on the respective moduli spaces of these theories then follow from this determination of τ, via $ds^2 = \mathrm{Im}(\tau)\, da d\bar{a}$.

We have given a brief review of Seiberg–Witten theory and its extension to include massive fundamental matter multiplets in this section. Using electric-magnetic duality, global symmetries and physical reasoning, and also properties of $\mathcal{N} = 2$ supersymmetry and holomorphy, the exact form of the low energy Wilsonian effective actions of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory and $SU(2)$ SQCD with $N_f \leq 4$ fundamental matter multiplets have been proposed by Seiberg and Witten. In Section 5.4 we will describe some of the work which followed the breakthrough results of Seiberg and Witten [170, 171]. These include the generalization of their analysis to larger and different gauge groups, and also to include different forms of matter. In Chapter 6 we will describe the tests of the proposed exact results using field theoretic methods, and the matching of the proposed exact results in $\mathcal{N} = 2$ supersymmetric gauge theories with instanton calculations.

Other Results in Seiberg–Witten Theory

We now briefly describe other work pertaining to the Seiberg–Witten solution for low energy $\mathcal{N} = 2$ $SU(2)$ SQCD and generalizations of it. There have been numerous comments made regarding Seiberg–Witten theory, and other results derived using it. In this subsection we also note other work which has been supplementary to Seiberg–Witten theory, and which explore issues involving the theory described in Section 5.3. A particularly remarkable development originating from Seiberg–Witten theory is its connection with the topology of four-manifolds. We remark briefly on this work below.

Of more immediate physical significance are the non-perturbative relations derived for Seiberg–Witten theory by Matone in [229], known as the Matone relation, and generalized to $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories in [232, 233, 234]. In the latter case, these have been used to derive recursion relations which enable one to calculate all of the coefficients $\mathcal{F}_k$ of the instanton contributions to the prepotential [234], in $SU(N)$ theories with and without matter multiplets.

The Matone relation is a renormalization group relation for the prepotential in $\mathcal{N} = 2$ $SU(2)$ Yang–Mills gauge theory [229]. The generalized Matone relation, derived in [232, 233, 234], relates the expansion coefficient of the k-instanton contribution to the $SU(N)$ prepotential, $\mathcal{F}_k|_{N,N_f}$, when there are N_f fundamental matter multiplets present, to the k^{th} power of the quantum modulus u_2 of the theory, as follows:

$$2\pi i\mathcal{F}_k|_{N=2,N_f} \;=\; u_2^k. \tag{5.201}$$

The Matone relation for $SU(2)$ has been tested by a two-instanton calculation in [230] and verified to all orders in k in [231]. The $SU(N)$ generalization of the Matone relation in [234] has also been verified to all orders in k in [224]. These all-orders tests represent the derivation of these non-perturbative relations from first principles using instanton calculus.

Another important result is the demonstration that the Seiberg–Witten solution of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory with and without fundamental matter contains more information than the exact series of instanton contribution expansion coefficients to the prepotential $\mathcal{F}$ [219]. The Seiberg–Witten solution provides non-trivial renormalization group information about $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory with N_f fundamental matter multiplets [219]. This is related to the renormalization group flow between the elliptic curves proposed for particular values of $N_f \leq 4$ described in Section 5.3.

Remarks have been made concerning the assumptions, both explicit and implicit, used by Seiberg and Witten [235, 237]. The uniqueness of the low energy Wilsonian effective Lagrangian in the Seiberg–Witten solution has also been demonstrated for both $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills and SQCD [238]. The electric-magnetic duality used in Seiberg–Witten theory has also been derived independently of the assumptions of

Seiberg–Witten theory [239].

We now briefly describe some of the mathematical work which has originated from Seiberg–Witten theory. The interest in Seiberg–Witten theory for pure mathematics originates with the work of Witten in [291]. In this work, Witten shows that a 'twisted' version (which has a precise mathematical definition) of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, to which the techniques of Seiberg–Witten theory can be applied, can be related to the topology of four-manifolds. The topology of four-manifolds [285], and in particular compact four-manifolds with boundary, was revolutionized one decade before Seiberg–Witten theory by Donaldson [286]. Donaldson used the study of $SU(2)$ Yang–Mills instantons on compact four-manifolds to define new topological invariants, known as Donaldson invariants. Donaldson invariants can be used to distinguish diffeomorphic four-manifolds, under certain conditions. These invariants complement the earlier work of Freedman [285], which can be used to distinguish homeomorphic four-manifolds. Unlike previous studies of four-manifolds, however, the use of gauge field theory, and in particular Yang–Mills instantons, now described in purely mathematical terms and divorced from their physical meaning, to study the topology of four-manifolds, was a completely new concept. The input of theoretical physics to the study of four-manifolds is an example of physics finding unexpected application to problems in pure mathematics. Unfortunately, the calculation of Donaldson invariants is especially difficult, and using the original Donaldson theory, the invariants have only been determined explicitly for special classes of four-manifolds. For further details on the Donaldson invariants and the application of instantons to four-manifold topology, we refer the reader to the reviews [288, 290]. Further mathematical background can be found in [289].

Witten's work in [291] enables one to define Donaldson invariants in terms of correlations functions of the twisted $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory. This greatly simplifies the calculation of Donaldson invariants for a wide class of four-manifolds (under certain conditions). This is a remarkable instance of theoretical physics finding use in the simplification of a mathematical problem. Using the duality of Seiberg–Witten theory, Witten was also able to define a dual twisted $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, which leads to new topological invariants in analogy with the Donaldson invariants. These new invariants were found by considering the analogue of the self-dual Yang–Mills field equations in the dual twisted $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, which are known as the monopole equations. The new topological invariants are known as the Seiberg–Witten invariants and can also be used to distinguish classes of four-manifolds under specific conditions. These new invariants essentially derive from theoretical physics and were not known previously. It is to be noted, however, that the Seiberg–Witten method for calculating Donaldson invariants, and the Seiberg–Witten invariants themselves, have yet to be put on a rigorous mathematical basis [297].

Detailed descriptions of the geometry of twisted $\mathcal{N} = 2$ supersymmetric gauge theories and its connections with the original physics of Seiberg–Witten theory are given in [292, 293]. Mathematical works which describe gauge field theory, Seiberg–Witten theory, the monopole equations and the relation of these to four-manifolds can be found in [294, 295, 296, 298, 299]. We note that the review [299] is particularly detailed and contains a comprehensive list of mathematics references.

5.4 Generalizations of Seiberg–Witten Theory

In this section we will describe some of the subsequent work in $\mathcal{N} = 2$ supersymmetric gauge theories which make use of, or further generalize, the exact results proposed for the particular case of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory. Much of this section is devoted to the generalization of the proposed results of Seiberg–Witten theory [170, 171] to different gauge groups of general rank. We will focus on $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories, which are the primary motivation and subject of this thesis. In this section we follow the reviews [190, 191] and the original papers [173, 174, 175, 176, 177, 178, 179, 180, 181, 183, 184] regarding the generalization of Seiberg–Witten theory to gauge group $SU(N)$. Other work related to generalizing Seiberg–Witten theory include [185, 186, 187] and those works cited in Subsection 5.4.2 below.

One initial generalization of Seiberg–Witten theory appears in the first paper of Seiberg and Witten [170]. This is the generalization of the formalism for the metric on the moduli space to spacetime dimensions greater than four. Lower dimensional analogues of

Seiberg–Witten theory are suggested to already exist in the field of integrable systems. A starting point for the extension of the analysis of Seiberg and Witten to gauge groups other than $SU(2)$ is the parameterization of the classical moduli space. In Section 5.3, it was described how the correct parameterization of the $SU(2)$ moduli space is given in terms of functions of the scalar field vacuum expectation value a which are invariant under Weyl group reflections. Such Weyl invariants are given by the characteristic equation:

$$\det(\lambda - \phi) = 0, \tag{5.202}$$

where ϕ is the scalar field component of the chiral superfield multiplet Φ in the theory. For $SU(2)$, classically one has $\phi = \frac{1}{2}a\sigma_3$. To obtain Weyl invariant functions of ϕ, one can expand Eq. (5.202) in powers of λ, and the coefficients of the resulting polynomial in λ will be Weyl invariant functions of ϕ. The generic roots of this polynomial are labelled a_i, $i = 1, \ldots, N$, where N is the rank of the gauge group. The roots a_i are the eigenvalues of ϕ. The polynomial in λ given by expanding the characteristic equation Eq. (5.202) has the form:

$$\lambda^N + \lambda^{N-2} \sum_{i<j} a_i a_j - \lambda^{N-3} \sum_{i<j<k} a_i a_j a_k + \cdots + (-1)^N \prod_{i=1}^{N} a_i = 0. \tag{5.203}$$

We shall first describe the proposed generalizations of Seiberg–Witten theory to $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories, with and without fundamental matter multiplets, in Subsection 5.4.1. Then we briefly describe the generalizations to other gauge groups and products of gauge groups in Subsection 5.4.2. In Subsection 5.4.2, we also mention the generalizations of Seiberg–Witten theory coupled to types of matter other than matter transforming under the fundamental representation of the gauge group.

We note that the Seiberg–Witten elliptic curve is not unique. Other elliptic curves may describe the moduli space of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory in an equally accurate way. When Seiberg–Witten theory is generalized to other gauge groups there is the possibility that several curves exist which appear to accurately describe the relevant moduli space of vacua. This is the scenario in the case of the generalization to $SU(N)$, which we shall describe in Subsection 5.4.1.

5.4.1 $\mathcal{N} = 2$ Supersymmetric $SU(N)$ Gauge Theories

After the defining papers of Seiberg–Witten theory [170, 171], theorists applied the techniques in these papers to other supersymmetric gauge theories, to propose analogous exact results. One such important application was the generalization of the work of Seiberg and Witten to $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories, for any $N > 1$. Given that many phenomenological non-Abelian field theories involve special unitary gauge groups, for example $SU(3)$ gauge symmetry in QCD, to propose exact results in these theories may prove useful for the study of non-supersymmetric non-Abelian gauge theores. Indeed, the gauge symmetries which appear to be realized exactly in nature are those which belong to the $SU(N)$ group, or Abelian subgroups of it, such as $U(1)$.

The moduli space of $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories can be parameterized using Weyl invariant functions of the scalar field vacuum expectation values in the theory. To obtain these functions, the expansion of the characteristic equation given in Eq. (5.203) can be used. When the gauge group is $SU(N)$, there are N vacuum expectation values a_i. The scalar field ϕ in the theory will be an $N \times N$ dimensional complex matrix with diagonal entries given by a_i:

$$\phi = \mathrm{diag}(a_1, a_2, \ldots, a_N). \tag{5.204}$$

In addition to unitarity, ϕ must obey the following condition in order to transform as an element of $SU(N)$:

$$\mathrm{Tr}(\phi) = \sum_i a_i = 0, \tag{5.205}$$

which ensures that the matrix ϕ has a unit determinant. The Weyl invariant functions obtained from the expansion of the characteristic equation have the form of symmetric polynomials $c_j(\phi)$ in terms of the variables a_i [190]:

$$u_i(\phi) = \sum_{n_1<n_2<\cdots<n_i}^{N} a_{n_1} a_{n_2} \cdots a_{n_i}. \tag{5.206}$$

The formula Eq. (5.206) can be illustrated for $N = 3$, for which the Weyl invariant functions which parameterize the $SU(3)$ moduli space are as follows:

$$0 = a_1 + a_2 + a_3, \tag{5.207}$$

$$u_1 = \frac{1}{2}\mathrm{Tr}(\phi^2) = -(a_1a_2 + a_1a_3 + a_2a_3) = a_1^2 + a_2^2 + a_1a_2, \tag{5.208}$$

$$u_2 = -\frac{1}{3}\mathrm{Tr}(\phi^3) = a_1a_2a_3 = a_1a_2(a_1 + a_2). \tag{5.209}$$

In general, there will be $N - 1$ such parameters for the $SU(N)$ moduli space, which we denote by u_i. These are the moduli of the classical moduli space. The variables u_i are not equal to the eigenvalues a_i of the adjoint scalar field ϕ, since these do not provide a correct parameterization of the moduli space, but are Weyl (and gauge) invariant functions constructed from the set $\{a_i\}$. For $N = 2$, there is only one parameter, u, which parameterizes the classical $SU(2)$ moduli space.

In this theory, the variables which correspond to the vacuum expectation value a of the scalar field component of the vector multiplet and its dual a_D are a_{ik} and a_{Djk}, where $i, j, k = 1, \ldots, N$. The variables a_{ij} and a_{Dij} are matrices. The complexified coupling constant τ_{ij} in these theories is also a matrix, and is given by:

$$\tau_{ij} = (a^{-1})_{ik} a^k_{Dj}. \tag{5.210}$$

The hyperelliptic curves proposed to describe the moduli space will correspond to some form of complex manifold. Such a surface is required to possess $2N$ periods, there being N each for each of the variables a_{Di} and a_i. The period matrix τ'_{ij} of the surface must also have a complex part which is positive definite, $\mathrm{Im}(\tau'_{ij}) > 0$, and which is $Sp(2N-2, \mathbb{Z})$ invariant. A candidate surface with these properties is a genus $N - 1$ Riemann surface, which is described by hyperelliptic curves of degree $2N$. In general, these complex curves have the form:

$$y^2 = \prod_{j=1}^{2N} (x - e_j), \tag{5.211}$$

where $\{e_j\}$ is the set of roots of the curve. At each of the $2N$ points $x = e_j$, the double cover of the Riemann surface has a branch point. Therefore, the double cover is a Riemann sphere.

Through the proposed identification of the moduli space of vacua of the theory with the moduli space of a genus $N-1$ Riemann surface, the complexified coupling matrix τ_{ij} is identified with the period matrix τ'_{ij}, so that $\tau_{ij} = \tau'_{ij}$. Then the variables a_{ik} and a_{Djk} can be calculated as the contour integrals of appropriate differential one-forms ω_k about a canonical homology basis of one-cycles γ_i on the Riemann surface, as in Seiberg–Witten theory:

$$a_{ik} = \oint_{\gamma_i} \omega_k, \quad a_{Djk} = \oint_{\gamma_j} \omega_k. \tag{5.212}$$

An inductive argument permits one to check the monodromies of the singularities for one value of $N > 2$ (since the Seiberg–Witten result is already assumed to be correct, $N \neq 2$ must be chosen) and then extend this result to $SU(N)$ for general $N > 2$. The most simple choice for these tests is the $N = 3$ curve [173, 174].

Hyperelliptic curves which purportedly describe the quantum moduli space of $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory have been proposed and studied in [173, 174, 178]. Hyperelliptic curves for the moduli spaces of $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD with N_f fundamental matter multiplets have also been proposed. These include the cases: $N_f \leq N$, [181]; $N_f \leq 2N$, [182, 184, 186]; $N = 3$ and $N_f = 6$, [183], and $N_f = 2N$, [185].

In this section we shall not describe the individual derivations of the elliptic curves proposed to describe the moduli space of $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with and without fundamental matter multiplets, but rather we shall outline the general techniques used to obtain them. We exhibit a particular elliptic curve for the $\mathcal{N} = 2$ $SU(N)$ theory as generic of the other curves proposed.

$\mathcal{N} = 2$ Supersymmetric $SU(N)$ Yang-Mills Gauge Theory

The pure $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory is given by the Lagrangian Eq. (3.86) of Section 3.4 in Chapter 3 with gauge group $SU(N)$. There have been many proposals for the exact low energy Wilsonian effective action of this theory, making use of the techniques introduced by Seiberg and Witten [170, 171]. In Seiberg–Witten theory, where the gauge group is $SU(2)$, the family of curves which are proposed to describe the quantum moduli space of the theory are elliptic curves which correspond to Riemann surfaces of genus one. The elliptic curves proposed for larger gauge groups correspond to Riemann surfaces of genus greater than one. Such elliptic curves are generically known as hyperelliptic curves due to this property.

Here we describe those hyperelliptic curves proposed to reproduce the moduli space of vacua of the pure $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory, by describing their generic properties with reference to a particular proposed hyperelliptic curve. The same hyperelliptic curve can be generalized to describe the same theory coupled to N_f fundamental matter multiplets, that is, $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD. There is overlap between the two cases since when there are $N_f = 0$ matter multiplets present, the hyperelliptic curve is expected to reproduce the moduli space of the pure $\mathcal{N} = 2$ $SU(N)$ Yang–Mills theory. Since the proposed hyperelliptic curves for $\mathcal{N} = 2$ $SU(N)$ SQCD include the $\mathcal{N} = 2$ $SU(N)$ Yang–Mills hyperelliptic curves as a special case, we shall not describe them here.

The hyperelliptic curves for these theories cannot be derived from first principles, just as the Seiberg–Witten elliptic curve was not derived from established physical techniques. Many attempts to obtain the hyperelliptic curves for the $\mathcal{N} = 2$ $SU(N)$ Yang–Mills gauge theory involve consistency requirements. One such requirement is the behaviour at large values of the moduli of the moduli space, that is, in the limit $|u_i| \to \infty$ ('infinity' on the moduli space), at which the gauge coupling becomes weak. The monodromies at these limits on the moduli space provide a non-trivial consistency condition on the hyperelliptic curve. Other monodromies on the moduli space, at singularities interpreted as points where monopoles and dyons become massless, which can be calculated from the hyperelliptic curve, will further permit checks that the choice of curve is correct or not. Other requirements are the correct asymptotic behaviour in the classical limit and scaling limits, R-symmetry, and compatibility with the BPS mass formula (given in Eq. (5.14)). The hyperelliptic curves proposed to describe the moduli space of $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory include those given in [173, 174, 178]. Many of the proposed hyperelliptic curves for the $SU(N)$ theory are variations on the curve presented by Klemm et. al [173]:

$$y^2 = \left(x^N - \sum_{i=2}^{N} u_i x^{N-1}\right)^2 - \Lambda^{2N}, \tag{5.213}$$

in which $\{u_i\}$ are the quantum moduli associated with the curve and Λ is the dynamical scale of the theory.

Once a hyperelliptic curve has been proposed for the $\mathcal{N} = 2$ $SU(N)$ theory, the procedure for calculating the period matrix τ_{ij} and the variables a_{ik}, a_{Djk}, as described above, can be applied. Given the matrices of variables $a_{ik}(u_i)$ and $a_{Djk}(u_i)$ in terms of the moduli u_i, the prepotential $\mathcal{F}$ of the low energy effective action of the theory can be calculated.

$\mathcal{N} = 2$ Supersymmetric $SU(N)$ QCD with N_f Matter Multiplets

When there are N_f fundamental matter multiplets, a similar set of requirements can be imposed in order to propose a suitable hyperelliptic curve for these theories. Hyperelliptic curves corresponding to Riemann surfaces of genus $N-1$ proposed for the $\mathcal{N} = 2$ $SU(N)$ Yang–Mills gauge theories can be used as starting points for the curves for $SU(N)$ SQCD with N_f matter multiplets.

The arguments used to deduce these curves are similar to those employed by Seiberg and Witten in their analysis of $\mathcal{N} = 2$ $SU(2)$ SQCD with N_f fundamental matter multiplets [171]. These include renormalization group flow, symmetries of the theory for special values of the multiplet masses, classical limits, and R-symmetry arguments. Hyperelliptic curves have been proposed for the following ranges of N_f: $N_f \leq 2N$ [181, 182]; $N_f = 2N$ [181, 183]. There has also been a hyperelliptic curve proposed for the $SU(N)$ theory when $N < 6$ with N_f matter multiplets [271]. All of these proposed hyperelliptic curves involve modular forms in the coefficients of the curve. Unfortunately, all of the hyperelliptic curves proposed for the $SU(N)$ with $N_f = 2N$ matter multiplets are not explicitly equivalent to each other [219]. This is related to the fact that the moduli spaces of these theories are not described by unique elliptic curves.

A unified treatment of the scale invariant $SU(N)$ theories, in which $N_f = 2N$, has been proposed by Argyres and Pelland [188]. Using the most general non-perturbative reparameterizations (or redefinitions) of the parameters within the hyperelliptic curves permitted, they show that the proposed hyperelliptic curves are special cases of a more general curve. The curve they propose for the $SU(N)$ theory with $N_f = 2N$ massive fundamental matter multiplets is:

$$y^2 = \left(x^N - \sum_{i=1}^{N-1} \tilde{u}_i x^{N-i-1}\right)^2 - \tilde{q} \prod_{r=1}^{2N} \left(x + \frac{1}{\sqrt{2}}\tilde{m}_r\right), \tag{5.214}$$

where the quantities $\{\tilde{u}_i, \tilde{q}, \tilde{m}_r\}$ are, respectively, the moduli, exponentiated (dimensionless) complexified coupling constant, and multiplet masses, each subject to non-perturbative reparameterization. We shall describe the matching of this proposed general hyperelliptic curve for the $SU(N)$ theory with $N_f = 2N$ massless matter multiplets to the results of first principles instanton calculations in Section 6.5 of Chapter 6.

5.4.2 Other Gauge Groups and Matter Content

We now briefly describe other exact results proposed for supersymmetric gauge theories with gauge groups other than $SU(N)$. These include the other classical gauge groups $SO(N)$ and $Sp(N)$. The exceptional groups G, F and A are also gauge groups, and we include a brief report of works which propose elliptic curves for $\mathcal{N} = 2$ supersymmetric Yang–Mills theory (with and without matter) with these gauge groups.

The hyperelliptic curves proposed for these theories have been derived using similar methods for the $SU(N)$ gauge group. The curves for both the $SO(N)$ and $Sp(N)$ gauge theories can also be derived by imposing conditions on the $SU(N)$ curves. Accordingly, all of these hyperelliptic curves involve coefficients which are modular forms.

$\mathcal{N} = 2$ Supersymmetric $SO(N)$ Gauge Theories

The hyperelliptic curves proposed for $\mathcal{N} = 2$ supersymmetric $SO(N)$ gauge theories coupled to differing numbers of N_f fundamental matter multiplets include those given for: $SO(2r+1)$ (or $SO(2N)$) with $N_f = 0$ [258, 265]; $SO(2r+1)$ with $N_f = 2r-1$ [259]; $SO(2r)$ with $N_f = 2r-2$ [260]; $SO(N)$ with $N_f = N-2$ [261], and $SO(2N)$, $N < 5$, with $N_f \geq 0$ [271].

$\mathcal{N} = 2$ Supersymmetric $Sp(N)$ Gauge Theories

The hyperelliptic curves proposed for $\mathcal{N}=2$ supersymmetric $Sp(N)$ gauge theories include those given for: $Sp(2r)$ with $N_f = 2r+2$ fundamental matter multiplets [260] and $USp(N)$ (the unitary symplectic groups) [260].

$\mathcal{N}=2$ Supersymmetric Gauge Theories with Exceptional Gauge Groups

The hyperelliptic curves proposed for the case of $\mathcal{N}=2$ supersymmetric gauge theories with exceptional gauge groups include those given for: G_2 and F_4 pure gauge theories [262]; E_6 pure gauge theory on a submanifold of the moduli space [262]; E_6 gauge theory with N_f matter multiplets [271]; E_7 gauge theory with N_f matter multiplets [271], and G_2 pure gauge theory [268, 266].
General hyperelliptic curves have also been claimed to be determined for all Lie gauge groups [267] and arbitrary classical gauge groups [269].

$\mathcal{N}=2$ Supersymmetric Gauge Theories with Product Gauge Groups

In the case of $\mathcal{N}=2$ supersymmetric gauge theories with gauge groups which are products of classical gauge groups, hyperelliptic curves include those proposed for scale invariant $SU(2)\times SU(2)$ pure gauge theory and scale invariant gauge theories with other product groups [270]. There have also been M-theoretic and brane theoretic derivations of hyperelliptic curves for $\mathcal{N}=2$ supersymmetric gauge theories with product groups [272].

Other Matter Content

Upon generalizing Seiberg–Witten theory to include the cases of other gauge groups, both classical and exceptional, and also product gauge groups, one direction in which to generalize further is that of coupling $\mathcal{N}=2$ supersymmetric Yang–Mills gauge theories or $\mathcal{N}=2$ SQCD theories to other types of matter multiplets. These include antisymmetric and symmetric tensor matter multiplets, which have been studied in [271, 273] in the context of M-theory and brane theory.

Chapter 6

Instanton Tests of the Exact Results in $\mathcal{N} = 2$ Supersymmetric Gauge Theories

6.1 Introduction

The exact results proposed for the low energy Wilsonian effective actions of $\mathcal{N} = 2$ supersymmetric gauge theories with various gauge groups were described in Chapter 5. These results use the indirect techniques introduced by Seiberg and Witten for the $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory with and without fundamental matter multiplets. These techniques are unconventional and make use of many assumptions and a hypothesis connecting the moduli space of vacua of these theories with the moduli space of specific Riemann surfaces. The use of elliptic and hyperelliptic curves in the method to exactly derive the $\mathcal{N} = 2$ prepotentials which describe the low energy effective dynamics of these theories appears ad hoc. However, the non-perturbative contributions to the $\mathcal{N} = 2$ prepotential are purportedly calculated exactly in the Seiberg–Witten solution of the low energy dynamics of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory. Due to supersymmetry, these contributions arise only from instantons. One can calculate the instanton contributions from first principles using conventional field theoretic methods, namely semi-classical, or saddle-point, calculations about instanton configurations. In this chapter we briefly describe instanton calculus and its use in testing the proposed exact solutions of $\mathcal{N} = 2$ supersymmetric $SU(2)$ and $SU(N)$ gauge theories. A quantitative comparison of the results of instanton calculations with predictions from the proposed exact results will then validate, to some extent, the assumptions and hypotheses used in the construction of the low energy effective supersymmetric gauge theories. This comparison can also be used to fix the arbitrariness present in the proposed exact results, and thus ensure that the proposed exact results will be equivalent to those determined from field theory.

In Section 6.2 we describe the fundamental aspects of instanton calculus, and in particular focus on instanton calculus using instanton collective co-ordinates and ADHM instanton configurations, as described in Chapter 2. In Subsection 6.2.1 we describe the $\mathcal{N} = 2$ supersymmetric generalization of instanton calculus. We then describe the instanton tests of Seiberg–Witten theory so far performed in Section 6.3. These include various one-instanton tests and a two-instanton test. In Section 6.4 we describe the instanton tests which have been performed for the exact results proposed for $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories. Following these we describe the matching of proposed exact results to instanton predictions for Seiberg–Witten theory in Section 6.5. This involves fixing the non-perturbative parts of the proposed exact results to agree with instanton predictions through matching schemes. We describe in detail the matching of the proposed exact results in $\mathcal{N} = 2$ $SU(N)$ SQCD with one-instanton calculations in Subsection 6.5.1.

The examples of instanton calculus we give are illustrative rather than demonstrative. This is because the new results reported in Subsection 6.5.1 of this chapter regarding the matching of instanton calculations and proposed exact results make use of previous work derived from instanton calculus, but do not require explicit instanton calculations in the actual matching of results.

6.2 Instanton Calculus

The exact solutions of the self-dual field equations which minimize the classical Euclidean Yang–Mills gauge field action known as instanton configurations were described in detail in Chapter 2. In this section we briefly outline the use of instanton configurations in quantum field theory and in $\mathcal{N} = 2$ supersymmetric gauge theories in particular.

The following brief review of instanton calculus is based on results which have been reviewed in [63, 64, 224], which are based upon the pioneering work in [37, 40, 39], the canonical work in Osborn:1978rn,cftg,cgt, and the work on multi-instantons in [42, 43], and also [44, 45, 46, 45, 47, 48], and other results in [38].

In this section we will describe the calculational methods in quantum field theory which involve instantons, generically known as the instanton calculus. The fundamental techniques in field theory which use instantons, or other solutions of the classical field equations, are the semi-classical method and the collective co-ordinate method. Below we describe these methods in preparation for the instanton tests presented in Sections 6.3 and 6.4, which make use of the supersymmetric multi-instanton calculus described in Subsection 6.2.1.

The Semi-Classical Approximation

The semi-classical approximation is an approximation method which interpolates between the classical and quantum versions of a physical theory. Use is made of a known classical field configuration, the behaviour of which with respect to the theory's action is also known. The classical configuration is then perturbed by quantum fluctuations and the path integral is expanded in terms of these fluctuations. The dominant non-perturbative effects in field theory at weak coupling are instantons.

In the path integral formalism, the semi-classical method can be illustrated by considering the instanton contribution to the partition function $Z[J]$ of the theory. We now describe the particular case of the instanton contribution to the partition function of a real (bosonic, quantum) scalar field ϕ. A scalar field theory could only receive instanton contributions if it is coupled to a Yang–Mills (or non-Abelian) gauge field theory; we take this to be implicit in what follows. IN a later paragraph, we turn to the case of gauge fields only. In general, the partition function $Z[J]$ for a four dimensional field theory will be of the form:

$$Z[J] \;=\; N\int [d\phi]\,\exp\left(-\frac{1}{g^2}S[\phi]+\int d^4x\,J\phi\right), \tag{6.1}$$

where J are external sources, $S[\phi]$ is the Euclidean action which is real and bounded from below, and N is an infinite normalization factor. At weak coupling, that is, for small values of g^2, the path integral will be dominated by field configurations which locally minimize the Euclidean action. The most simple configurations which do this are the minima of the classical potential of the theory; expanding about these configurations gives standard perturbation theory, or the loop expansion. Other such minimizing configurations cannot be dealt with perturbation theory, however. There exist inherently non-perturbative effects which will give a finite action of the theory, and these are instantons. The term 'semi-classical' refers to the factor of $1/g^2$ multiplying $S[\phi]$ in Eq. (6.1): if powers of h are restored, then there is a factor of $1/h^2$ also multiplying $S[\phi]$, and thus the classical limit $h \to 0$ can be identified with the weak coupling limit $g \to 0$.

If a generic quantum field theory with generating functional as in Eq. (6.1) exists with no external sources present, then $J = 0$, and the path integral in Eq. (6.1) becomes:

$$Z[0] \;=\; \int [d\phi]\,\exp\left(-\frac{1}{g^2}S[\phi]\right). \tag{6.2}$$

Let the configuration of fields which specify an instanton solution be $\phi_{\rm cl}$. Then the instanton configuration is a local minimum of the Euclidean action $S[\phi]$ which obeys the Euler–Lagrange equation:

$$\left.\frac{\delta S}{\delta\phi}\right|_{\phi=\phi_{\rm cl}} \;=\; 0. \tag{6.3}$$

To approximate the quantum field ϕ semi-classically, let the quantum fluctuations about the instanton configuration $\phi_{\rm cl}$ be $\phi_{\rm qu}$, so that ϕ may be written as:

$$\phi \;=\; \phi_{\rm cl}+\phi_{\rm qu}. \tag{6.4}$$

A Taylor expansion of the action $S[\phi]$ about the field ϕ may then be performed in the background of the instanton configuration $\phi_{\rm cl}$:

$$S[\phi] \;=\; S\left[\phi_{\rm cl}\right]+\frac{1}{2}\int d^4x\,\phi_{\rm qu}(x)\widehat{M}\phi_{\rm qu}(x)+\mathcal{O}(\phi_{\rm qu}^3). \tag{6.5}$$

The expansion has only been continued to second order in the field $\phi_{\rm qu}$ as this corresponds to a one-loop approximation, which throughout our description of the instanton calculus is the required approximation. Furthermore, the linear term $(\delta S/\delta\phi)$ is absent in Eq. (6.5) due to Eq. (6.3). The factor $\widehat{M}$ inside the integral in Eq. (6.5) is an operator valued quantity given by:

$$\widehat{M} \;=\; \left.\frac{\delta^2 S}{\delta\phi^2}\right|_{\phi\;=\;\phi_{\rm cl}} . \tag{6.6}$$

In general the operator $\widehat{M}$ will have a complete orthonormal set of eigenfunctions $\{\phi_i\}$ and associated eigenvalues $\{\epsilon_i\}$. The eigenfunctions $\{\phi_i\}$ can be used as a basis for the quantum fluctuations in ϕ:

$$\phi_{\rm qu} \;=\; \sum_i c_i\phi_i. \tag{6.7}$$

This permits one to write the second order term in Eq. (6.5) as:

$$\frac{1}{2}\int d^4x\,\phi_{\rm qu}(x)\widehat{M}\phi_{\rm qu}(x) \;=\; \frac{1}{2}\sum_i c_i^2||\phi_i||^2\epsilon_i, \tag{6.8}$$

where we have defined:

$$||\phi_i||^2 \;=\; \int d^4x\,\phi_i(x)\cdot\phi_i(x) \tag{6.9}$$

which is the L^2 norm for ϕ_i.

If the integration in this approximation is to be performed, then the functional integration measure $[d\phi]$ in Eq. (6.5) must be expressed in terms of the basis coefficients $\{c_i\}$. The appropriate integration measure is defined to be:

$$[d\phi] \;=\; \prod_i \frac{||\phi_i||}{\sqrt{2}}dc_i. \tag{6.10}$$

When the expressions in Eqs. (6.8,6.9,6.10) are substituted into the semi-classical form of the action in Eq. (6.5), the functional integration for $Z[0]$ in Eq. (6.2) can be performed since the exponential in the integrand in Eq. (6.2) is only quadratic in $\{c_i\}$. The functional integration therefore reduces to a Gaussian integral in $\{c_i\}$ and gives the one loop instanton contribution to the partition function $Z[0]$ as:

$$Z^1[0] \;=\; N(\det\widehat{M})^{-\frac{1}{2}}\exp\left(-S\left[\phi_{\rm cl}\right]\right), \tag{6.11}$$

in which:

$$\widehat{\rm M} \;=\; \prod_i \epsilon_i. \tag{6.12}$$

This is a simplified calculation because an implicit assumption has been made regarding the eigenvalues $\{\epsilon_i\}$ of $\widehat{M}$. A generic operator $\widehat{M}$ will possess zero-valued eigenvalues. If at least one single eigenvalue ϵ_i is zero, then the determinant of $\widehat{M}$ in Eq. (6.12) vanishes. Then it follows that the expression for $Z^1[0]$ in Eq. (6.11) is divergent and thus ill-defined. This indicates that a more general method of calculating the instanton contribution to the partition function $Z[0]$ is required if the calculations using the semi-classical approximation in a generic instanton background are to be made. Fortunately, a method known as the collective co-ordinate method exists which permits this.

The Collective Co-ordinate Method

The collective co-ordinates of the instanton configuration are directly related to the symmetries of the classical theory which the instanton solution breaks. Collective co-ordinates for the BPST instanton and ADHM multi-instantons were described in Chapter 2. The broken classical symmetries of the theory are directly related to the presence of zero-eigenvalues of the operator $\widehat{M}$. If the instanton configuration $\phi_{\rm cl}$ breaks n classical symmetry generators, then it will be parameterized by n collective co-ordinates $\{\mathfrak{X}_i\}$, $i = 1, \ldots, n$. The set of instanton configurations $\phi_{\rm cl}$ then consists of an n-dimensional region of field configuration space in which the action $S[\phi]$ has a constant minimum value. Since the space has an n-dimensional basis, there exist n independent directions along which the action $S[\phi]$ is constant. Along these directions the action must not vary with ϕ and therefore in these directions the eigenvalues of the operator $\widehat{M}$ must be zero. The set of zero-eigenvalues of $\widehat{M}$ are associated with a set of zero eigenfunctions of $\widehat{M}$ which are known as 'zero modes.' In a gauge field theory, these zero modes are subject to a gauge fixing condition and therefore the definition of these eigenfunctions as tangent vectors along the n directions of constant action is incomplete.

To illustrate how the collective co-ordinate method includes zero modes, let the zero modes of the operator $\widehat{M}$ be the first n eigenfunctions given by ϕ_i, where $i = 1, \ldots, n$. The eigenvalues of the zero modes vanish, and therefore the associated basis coefficients c_i, $i = 1, \ldots, n$ do not appear in the sum in Eq. (6.8) for the quadratic ϕ_i term. The consequence of this is that the integration over the basis coefficients c_i is divergent, as there are basis coefficients in the integration in which only zero eigenvalues are included. To circumvent this divergence, one can effect a change of integration variables. If one changes the n integration variables from the basis coefficients c_i to that of the instanton collective co-ordinates X_i, this is the collective co-ordinate method for performing the integration. The result is a convergent integration.

One can make the required change of variables by inserting a factor of unity into the path integral for $Z^1[0]$, via:

$$1 = \int d\mathfrak{X}_1 d\mathfrak{X}_2 \cdots d\mathfrak{X}_n (\det\Delta) \prod_{i=1}^{n} \delta(\phi - \phi_{\rm cl}, \phi_i), \tag{6.13}$$

where Δ is an $n \times n$ matrix given by:

$$\Delta_{ij} = \left(\frac{\partial \phi_{\rm cl}}{\partial \mathfrak{X}_i}, \phi_j\right) + \mathcal{O}(\phi - \phi_{\rm cl}) = \int d^4x \, \frac{\partial \phi_{\rm cl}(x)}{\partial \mathfrak{X}_i} \phi_j(x) + \mathcal{O}(\phi - \phi_{\rm cl}), \tag{6.14}$$

in which the $\mathcal{O}(\phi - \phi_{\rm cl})$ are neglected in the one loop approximation, consistent with the expansion in Eq. (6.5). The semi-classical approximation can now be made for the field ϕ, and upon substituting Eq. (6.13) into Eq. (6.5) and integrating out the basis coefficients c_i, $i > n$ which are associated with zero modes, the one loop instanton contribution to the partition function $Z[0]$ can be written as:

$$Z^1[0] = N \int \left(\prod_{i=1}^{n} d\mathfrak{X}_i\right) \left(\prod_{i=1}^{n} dc_i\right) (\det\Delta) \left(\prod_{i=1}^{n} \frac{||\phi_i||}{\sqrt{2}} \delta(c_i ||\phi_i||^2)\right) (\det{}'\widehat{M})^{-\frac{1}{2}} \exp(-S[\phi_{\rm cl}]), \tag{6.15}$$

where the adjusted determinant of $\widehat{M}$ is given by:

$$\det{}'\widehat{M} = \prod_{i>n} \epsilon_i, \tag{6.16}$$

which now excludes all of the zero eigenvalues $\{\epsilon_i\}$, $i = 1, \ldots, n$. The delta functions in the collective co-ordinate integral Eq. (6.15) saturate the integrations over the basis coefficients c_i, $i = 1, \ldots, n$, so that these integrations vanish from the integral. The result is that the integral expression for $Z^1[0]$ simplifies to:

$$Z^1[0] = N \int d\mathfrak{X}_1 d\mathfrak{X}_2 \cdots d\mathfrak{X}_n (\det\Delta) \left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \frac{1}{||\phi_i||}\right) (\det{}'\widehat{M})^{-\frac{1}{2}} \exp(-S[\phi_{\rm cl}]). \tag{6.17}$$

The integration over the collective co-ordinates in Eq. (6.17) is convergent, and the result for $Z^1[0]$ is now formally equivalent to the divergent quantity in Eq. (6.11). The instanton contribution to $Z[0]$ must also be correctly normalized. This can be acheived by setting the infinite normalization constant N to the value:

$$N = (\det\widehat{M}_0)^{\frac{1}{2}}, \tag{6.18}$$

where $\widehat{M}_0$ is the operator $\widehat{M}$ in the particular background $\phi_{\rm cl} = 0$, which corresponds to the trivial vacuum solution of the classical field equations.

The one loop instanton contribution to the partition function in Eq. (6.17) must also be renormalized, assuming that the quantum field theory under consideration is a renormalizable one. This is because the factor $(\det'\widehat{M})^{-\frac{1}{2}}$ will generically have one loop ultra violet divergences, arising from large eigenvalues of the operator $\widehat{M}$, of which there are an infinite number. Therefore, renormalization of this non-perturbative calculation must be performed.

A similar treatment holds for the (one loop) instanton contribution to the partition function for theories involving fields other than real scalar field ϕ and combinations of fields. For fermionic zero modes, the collective co-ordinates are Grassmann valued, and the calculation becomes more complicated.

Quantities other than the partition function $Z[0]$ in a Yang–Mills gauge field theory also receive instanton contributions. In general, Green's functions in the theory will receive instanton contributions. The collective co-ordinate method can be used in the one loop semi-classical approximation, and the insertion of unity into the path integral will have the basis coefficient integrations saturated by the classical instanton background.

The Collective Co-ordinate Integral in Gauge Theories

We now turn to the semi-classical approximation and the collective co-ordinate method in gauge field theories. Our description is necessarily schematic. We describe the collective co-ordinate integral in pure $SU(N)$ Yang–Mills gauge theory for definiteness. This is also an appropriate choice since the $\mathcal{N} = 2$ supersymmetric generalization of $SU(N)$ Yang–Mills gauge theory is the field theory of primary focus in this thesis.

Let A^a_m be the $SU(N)$ gauge field. If we suppress the gauge group indices, then this gauge field has the semi-classical expansion:

$$A_m = A_m^{\rm cl} + A_m^{\rm qu}, \tag{6.19}$$

where $A_m^{\rm cl}$ is the instanton configuration and $A_m^{\rm qu}$ are the quantum fluctuations about it. The gauge of the instanton gauge field $A_m^{\rm cl}$ must be fixed otherwise local gauge transformations may be included in the fluctuations of this field, and the value of the invariant action altered. A convenient gauge is the covariant background gauge, in which:

$$D_m^{\rm cl} A_m^{\rm qu} = 0, \tag{6.20}$$

where $D_m^{\rm cl}$ is the covariant derivative in the instanton background, which acts as $D_m^{\rm cl}(A_n^{\rm qu})^a = \partial_m (A_m^{\rm qu})^a + igf^{abc}(A_m^{\rm cl})^b (A_n^{\rm qu})^c$ This gauge imposes the condition that the quantum fluctuations $A_m^{\rm qu}$ are orthogonal to infinitesimal gauge transformations of the classical field $A_m^{\rm cl}$ in field space. An infinitesimal gauge transformation of $A_m^{\rm cl}$ is given by:

$$A_m^{\rm cl} \to A_m^{\rm cl\prime} = A_m^{\rm cl} + D_m^{\rm cl}\Lambda, \tag{6.21}$$

for some arbitrary function Λ. The orthogonality condition can then be expressed as:

$$\int d^4x \, {\rm Tr}\left(A_m^{\rm qu} D_m^{\rm cl}\Lambda\right). \tag{6.22}$$

Integration by parts and the arbitrariness of Λ can be used to show that the condition in Eq. (6.22) is equivalent to the covariant background gauge in Eq. (6.20). It can be further shown that the gauge field zero modes satisfy the covariant Weyl equation, and that each gauge zero mode can be interpreted as two independent solutions of the Weyl equation.

The classical Yang–Mills action for the $SU(N)$ gauge theory including Faddeev–Popov gauge-fixing and ghost terms is given by:

$$\begin{aligned} S[A,\bar{\eta},\eta] = \frac{1}{2}\int d^4x \, [&{\rm Tr}(v_{mn}v^{mn}) \\ &+C^2(A_m) + 2\mathcal{L}_{gh}(\bar{\eta},\eta)] \,, \end{aligned} \tag{6.23}$$

where, as previously, the non-Abelian gauge field strength v_{mn} is given by:

$$v_{mn}^a = \partial_m A_n^a - \partial_n A_m^a + igf^{abc}A_m^b A_n^c, \tag{6.24}$$

and $C(A_m)$ is the gauge-fixing term to which the Lagrangian $\mathcal{L}_{gh}(\bar{\eta},\eta)$ for the ghost fields $\{\bar{\eta},\eta\}$ is associated. The generic form of a functional path integral in a quantum gauge field theory is given by:

$$W = N\int[dA_m][d\bar{\eta}][d\eta]\exp(-S[A_m,\bar{\eta},\eta]). \tag{6.25}$$

which is analogous to the path integral given for a quantum scalar field theory in Eq. (6.1). The semi-classical approximation in Eq. (6.19) for the gauge field A_m can be substituted into $S[A]$, yielding, after some rearrangement:

$$\begin{aligned} S[A] =&\frac{8\pi^2k}{g^2}+\frac{1}{2}\int d^4x\mathrm{Tr}\left[2(D^{\mathrm{cl}}_mA^{\mathrm{qu}}_n)^2-2(D^{\mathrm{cl}}_mA^{\mathrm{qu}}_m)^2\right.\\ &\left.-4igv^{\mathrm{cl}}_{mn}[A^{\mathrm{qu}}_m,A^{\mathrm{qu}}_n]+C^2(A_m)+2\mathcal{L}_{\mathrm{gh}}(\bar{\eta},\eta)\right]+\mathcal{O}\left((A^{\mathrm{qu}}_n)^3\right). \end{aligned} \tag{6.26}$$

When the covariant background gauge is fixed according to Eq. (6.20), the gauge fixing term and the ghost Lagrangian in the Faddeev–Popov gauge fixing procedure then have the form:

$$\int d^4x\, C^2(A_m) = 2\int d^4x\,\mathrm{Tr}\left[(D^{\mathrm{cl}}_mA^{\mathrm{qu}}_m)^2\right], \tag{6.27}$$

$$\int d^4x\,\mathcal{L}_{\mathrm{gh}}(\bar{\eta},\eta) = -2\int d^4x\,\mathrm{Tr}\left[\bar{\eta}(D^{\mathrm{cl}})^2\eta\right]. \tag{6.28}$$

Upon substituting these terms into the action in Eq. (6.23), one obtains the following for the gauge-fixed one loop semi-classical expansion of $S[A]$:

$$S[A,\bar{\eta},\eta] = \frac{8\pi^2}{g^2}+\frac{1}{2}\int d^4x\,(A^{\mathrm{qu}}_m)^a(\Delta^{(+)})^{ab}_{mn}(A^{\mathrm{qu}}_m)^b+\int d^4x\,\bar{\eta}^a(\widehat{M}_{\mathrm{gh}})^{ab}\eta^b, \tag{6.29}$$

in which we have defined the operators:

$$(\Delta^{(+)})^{ab}_{mn} = -(D^{\mathrm{cl}})^2-2gf_{abc}(v^{\mathrm{cl}}_{mn})^b, \tag{6.30}$$

$$\widehat{M}_{\mathrm{gh}} = -(D^{\mathrm{cl}})^2\eta^a. \tag{6.31}$$

The operator $\Delta^{(+)}$ is a gauge field fluctuation operator which, along with its companion $\Delta^{(-)}$, are important in describing the fluctuations about the instanton configuration. The companion fluctuation operator $\Delta^{(-)}$ is defined by:

$$(\Delta^{(-)})^{ab}_{mn} = -(D^{\mathrm{cl}})^2-2gf_{abc}({}^*v^{\mathrm{cl}}_{mn})^b, \tag{6.32}$$

where $^*v_{mn}$ is the dual of the gauge field strength v_{mn}, defined in Chapter 2. In a pure instanton background, with no anti-instantons present, the operator $\Delta^{(-)}$ becomes:

$$(\Delta^{(-)})^{ab}_{mn} = -(D^{\rm cl})^2. \tag{6.33}$$

We note that in a pure instanton background, the operator $\Delta^{(+)}$ possesses $4Nk$ zero modes, corresponding to the $4Nk$ physical collective co-ordinates of an $SU(N)$ k-instanton. We denote these zero modes as $A^{(i)}_m$, $i = 1, \ldots, 4Nk$. The operator $\Delta^{(-)}$ becomes a positive definite operator in an pure instanton background and has no normalizable zero modes. The ghost field operator $\widehat{M}_{\rm gh}$ also has no zero modes.

In analogy with Eq. (6.4), the quantum fluctuations $A^{\rm qu}_m$ can be written in terms of the basis formed by the eigenfunctions of the fluctuations operator $\Delta^{(+)}$. Strictly, one can also include the non-zero modes in this expansion, and we denote the non-zero modes by $\tilde{A}^{(i)}_m$, $i = 1, \ldots, 4Nk$. The expansion of $A^{\rm qu}_m$ is then:

$$A^{\rm qu}_m = \sum_n \xi^n \delta_n A^{(i)}_m + \tilde{A}^{(i)}_m, \tag{6.34}$$

where $\{\xi^n\}$ is the set of expansion coefficients, in direct analogy to the scalar field collective co-ordinate method. The non-zero mode fluctuations $\tilde{A}^{(i)}_m$ are orthogonal to the zero modes in a functional sense. Expanding $A^{\rm qu}_m$ as in Eq. (6.34), the functional integration over the measure $[dA_m]$ in Eq. (6.25) can be written as:

$$\int [dA_m] = g^{-4Nk} \int [d\tilde{A}_m] \left\{ \sqrt{\det\, g(\mathfrak{X})}\, \prod_n \frac{d\xi^n}{\sqrt{2\pi}} \right\}. \tag{6.35}$$

Here, $g(\mathfrak{X})$ is the metric on the zero modes. The term in braces is the integral over the zero mode subspace. The factors of g multiplying the measure arise from the fact that we included a factor of g^2 in the definition of the metric, as:

$$g_{mn}(\mathfrak{X}) = -2g^2 \int d^4x\, {\rm tr}_N A^{\rm qu}_m(x) A^{\rm qu}_n(x). \tag{6.36}$$

The non-zero mode fluctuations $\tilde{A}_m$ and the ghost fields $\{\bar{\eta}, \eta\}$ can now be integrated out, which produces the following determinant factors:

$$\det(-(D^{\rm cl})^2) \left(\det{}' \left[\Delta^{(+)}\right]\right)^{-1}. \tag{6.37}$$

As is conventional, the prime on the determinant indicates that the operator $\Delta^{(+)}$ has zero modes; hence $\det\Delta^{(+)}$ is modified to $\det'\Delta^{(+)}$ as zero modes must be excluded in the product over eigenvalues that defines the (convergent) determinant. The leading order expression for the functional integral in the charge-k sector is then:

$$\frac{e^{2\pi i k\tau}}{g^{4Nk}}\int\left\{\sqrt{\det g(\mathfrak{X})}\prod_n\frac{d\xi^n}{\sqrt{2\pi}}\right\}\det(-(D^{\rm cl})^2)\left(\det'\left[\Delta^{(+)}\right]\right)^{-1}. \tag{6.38}$$

The theory must be regularized in order to remove the divergences present in the determinants of the fluctuations operators arising from zero modes. A suitable regularization scheme which simplifies instanton calculations is the Pauli–Villars scheme [58], described in connection with these determinants below. The advantages of using the Pauli–Villars scheme in supersymmetric gauge theories will be apparent in Subsection 6.2.1.

Using the collective co-ordinate method, an appropriate insertion of unity can be used to express the collective co-ordinate integral in Eq. (6.25) for a gauge field theory. This insertion of unity has the following form:

$$1\equiv\int\prod_\mu d\mathfrak{X}^\mu\left|\det\left(g_{\mu\nu}(\mathfrak{X})-2g^2\int d^4x\,{\rm tr}_N A_n^{\rm qu}\frac{\partial\delta_\nu A_n^{\rm qu}}{\partial\mathfrak{X}^\mu}\right)\right|\prod_\mu\delta\left(\sum_\mu\epsilon^\mu g_{\mu\nu}(\mathfrak{X})\right). \tag{6.39}$$

This can be substituted into Eq. (6.25) in order to saturate the integrations over the expansion coefficients ξ^n, which is achieved via the delta functions in Eq. (6.39). The result is that the leading order k-instanton contribution to the path integral in the weak coupling limit has the following form, which is referred to as the collective co-ordinate integral:

$$\int[dA_n]\,[db]\,[dc]\,e^{-S[A,b,c]}|_{\text{charge-}k}\ \overset{g\to0}{\longrightarrow}\ \frac{e^{2\pi ik\tau}}{g^{4Nk}}\int_{\mathfrak{M}_k}\omega\cdot\frac{\det(-\mathcal{D}^2)}{\det'\Delta^{(+)}}, \tag{6.40}$$

where $\mathcal{D}^2$ is the covariant Laplacian operator. In the collective co-ordinate integral the factor ω is the canonical volume form on the moduli space $\mathfrak{M}_k$, associated to the metric $g_{\mu\nu}(\mathfrak{X})$:

$$\int_{\mathfrak{M}_k}\omega\equiv\int\sqrt{\det g(\mathfrak{X})}\prod_\mu\frac{d\mathfrak{X}^\mu}{\sqrt{2\pi}}. \tag{6.41}$$

In the integrand of Eq.(6.40), the volume form ω is multiplied by a non-trivial function on $\mathfrak{M}_k$ equal to the product of the ratios of the determinants of the operators governing the Gaussian fluctuations of the gauge field and ghosts in the instanton background.

We note that when calculating a correlation function $\langle\mathcal{O}_1(x^1)\cdots\mathcal{O}_n(x^n)\rangle$ in the semi-classical approximation, the field insertions $\mathcal{O}_i(x^i)$ are to be replaced by their values in the instanton background. To leading order in the semi-classical limit, then, these insertions are functions of the collective co-ordinates.

To proceed with instanton calculations, an explicit expression for the volume form is required. Using the description of the instanton moduli space $\mathfrak{M}_k$ as a hyper-Kähler quotient of flat space, i.e. $\mathbb{R}^4$, which was briefly outlined in Chapter 2, the authors of [224] are able to construct the volume form for the $U(N)$ instanton moduli space. We do not describe this construction in detail, which is mathematically non-trivial. The ADHM constraints, which in this construction are termed moment maps, can be implemented in the volume form as Dirac delta functions. The vanishing of the moment maps, which define the instanton configurations, are written implicitly as part of this construction. Originally, this concept was part of an ansätz made for the supersymmetric instanton measure in [213, 214, 217, 222], which has proven highly successful in its application. This procedure is apparently the only plausible way in which to implement the ADHM constraints in the collective co-ordinate integrals of instanton calculus. Furthermore, integrations over the instanton moduli space must involve exact instanton configurations, otherwise the integration may be divergent. If the integrand of the integral over the instanton moduli space contains a special exact instanton solution, rather than the general exact instanton solution, for a specific value of k, there may exist flat directions in the integrand. When the integration over the instanton moduli space is performed, these flat directions may lead to divergences, so that the result is not convergent. Thus the most general exact k-instanton solutions must be used when calculating the contributions of these configurations to the path integral. Exact general instanton configurations may be obtained for small values of k using the ADHM construction as described in Chapter 2. However, this entails solving the k-instanton ADHM constraints in complete generality for a given value of k. The method of including the ADHM constraints as the arguments of Dirac delta functions gives rise to the possibility of performing integrations over general exact ADHM k-instanton configurations without explicitly solving the ADHM constraints. In general, this procedure is prohibitively difficult in comparison to solving the ADHM constraints, which at least can be explicitly and generally solved for $k \leq 2$. The volume form which we state here is that which has the residual $U(k)$ symmetry of the $U(N)$ ADHM construction unfixed. The volume form ω on the $U(N)$ k-instanton moduli space is then [224]:

$$\int_{\mathfrak{M}_k} \omega = \frac{C_k}{\mathrm{Vol}\, U(k)} \int d^{4k(N+k)}a \, |\mathrm{det}_{k^2}\mathbf{L}| \prod_{r=1}^{k^2} \prod_{c=1}^{3} \delta\left(\frac{1}{2}\mathrm{tr}_k T^r \left(\tau^{c\dot\alpha}_{\dot\beta} \bar a^{\dot\beta} a_{\dot\alpha}\right)\right), \qquad (6.42)$$

where $\mathbf{L}$ is a $k \times k$ Hermitian operator matrix dependent on the $U(N)$ ADHM submatrices, defined as:

$$\mathbf{L} \cdot \Omega = \tfrac{1}{2}\{\bar w^{\dot\alpha} w_{\dot\alpha}, \Omega\} + [a'_n, [a'_n, \Omega]], \qquad (6.43)$$

$$= \tfrac{1}{2}\{\bar w^{\dot\alpha} w_{\dot\alpha}, \Omega\} + \tfrac{1}{2}\bar a'^{\dot\alpha\alpha} a'_{\alpha\dot\alpha} - \bar a'^{\dot\alpha\alpha}\Omega a'_{\alpha\dot\alpha} + \tfrac{1}{2}\Omega \bar a'^{\dot\alpha\alpha} a'_{\alpha\dot\alpha}, \qquad (6.44)$$

where Ω is a generic $k \times k$ anti-Hermitian matrix of scalars upon which $\mathbf{L}$ acts. The following objects in the volume form Eq. (6.42) have also been defined:

$$\int d^{4k(N+k)}a \equiv \int \prod_{n=1}^{4} \prod_{r=1}^{k^2} d(a'_n)^r \prod_{i=1}^{k} \prod_{u=1}^{N} \prod_{\dot\alpha=1}^{2} d\bar w^{\dot\alpha}_{iu} dw_{\dot\alpha ui}, \qquad (6.45)$$

$$\mathrm{Vol}\, U(k) = \frac{2^k \pi^{k(k+1)/2}}{\prod_{i=1}^{k-1} i!}, \qquad (6.46)$$

$$C_k = 2^{-k(k-1)/2}(2\pi)^{2Nk}. \qquad (6.47)$$

The integrations over the ADHM submatrices a'_n in Eq. (6.45) and the ADHM constraints in the delta function in Eq. (6.42) are defined with respect to the generators of the residual symmetry group $U(k)$ in its fundamental representation. These generators are normalized such that $\mathrm{tr}_k T^r T^s = \delta^{rs}$. In Eq. (6.46), the volume of the $U(k)$ group is given, which is a constant. In Eq. (6.47), the normalization factor in the volume form Eq. (6.42), which can be found from the normalization of the metric on $\mathfrak{M}_k$, is given. Note that there is a different normalization constant C_k for each k. This expression for the volume form on the $U(N)$ instanton moduli space reproduces known results, such as the one-instanton $U(N)$ volume form [40].

With the general form of the volume form on the $U(N)$ k-instanton instanton moduli space established, the leading order k-instanton contribution to the path integral can be written explicitly. The result is the collective co-ordinate integral on the general instanton moduli space $\mathfrak{M}_k$, which follows from the weak coupling limit of the path integral given in Eq. (6.40). The instanton measure, which comprises the integrations over the volume form ω and the fluctuation determinant factors in Eq. (6.40) on $\mathfrak{M}_k$, exhibits an important property known as clustering or the cluster decomposition. This is a physical property of the instanton measure, which corresponds to the dilute instanton gas limit described in Chapter 2. All correct instanton measures should possess this property. In the most simple clustering limit, the volume form ω for a k-instanton configuration can be interpreted as well separated k_1-instanton and k_2-instanton configurations in specific regions of the instanton moduli space, as described in Chapter 2. In these regions, one expects that the instanton moduli space $\mathfrak{M}_k$ is approximately equal to $\mathfrak{M}_{k_1} \times \mathfrak{M}_{k_2}$. In this case the k-instanton volume form ω factorizes as follows:

$$\int_{\mathfrak{M}_k} \omega \ \rightarrow \ \frac{k_1!k_2!}{k!} \int_{\mathrm{M}_{k_1}} \omega \times \int_{\mathrm{M}_{k_2}} \omega. \tag{6.48}$$

In the completely clustered limit, in which the k-instanton configuration decomposes into a sum of k 1-instantons, the volume form decomposes as follows, using the normalization in Eq. (6.47):

$$\int_{\mathfrak{M}_k} \omega \ \rightarrow \ \frac{1}{k!} \int_{\mathfrak{M}_1} \omega \times \cdots \times \int_{\mathfrak{M}_1} \omega, \tag{6.49}$$

where the number of 1-instanton integrations in the decomposition on the right hand side is equal to k. The primary use of the cluster decomposition in instanton calculus is to check the validity of the proposed instanton measure. It also provides a check on the set of normalization constants C_k which appear in the volume form ω in Eq. (6.42).

We now turn to the fluctuation determinants which remain in the instanton measure in Eq. (6.40). In general, their evaluation is a highly non-trivial task. The determinants of the fluctuation operators were first evaluated in [18] in the one-instanton background. Further pioneering work was also carried out in [47]. However, these factors have not yet been expressed as functions of the instanton moduli space. Instead, they remain as implicit results dependent on spacetime integrals. Despite this, the form of the fluctuations operators determinant factors can be calculated in the ADHM k-instanton background. We do not state this result or the details of its derivation, but shall briefly describe it for completeness.

In the Pauli–Villars regularization scheme [58], one can express the fluctuation determinant factors in terms of the Pauli–Villars scale μ, which is the mass of the Pauli–Villars regulator fields, taken to be large [58]. Using this regularization scheme enables one to make definitions of the determinants which are properly regularized in the UV (high energy) region. The determinant factor in Eq. (6.40) can then be written as:

$$\frac{\det(-\mathcal{D}^2)}{\det'\Delta^{(+)}} = \mu^{4Nk}\frac{1}{\det(-\mathcal{D}^2)}, \tag{6.50}$$

so that the problem is now one involving the determinant of the covariant Laplacian operator $\mathcal{D}^2$. The formula Eq. (6.50) is derived using expressions for the Pauli–Villars regularized determinants $\det'\Delta^{(+)}$ and $\det\Delta^{(-)}$, for which $\det\Delta^{(-)} = [\det(-\mathcal{D}^2)]^2$. The problem of evaluating the fluctuations operators determinant factor is now reduced to one requiring the fluctuation determinant of a scalar field, upon which $-\mathcal{D}^2$ operates, existing in the adjoint representation of the gauge group. A series of results from previous work on instantons can be used to calculate this quantity. Following [224], one can use the fluctuation determinant of a scalar field transforming in the fundamental representation. Then, using a formula derived in [43] which relates this to the determinant for a scalar field transforming in the adjoint representation, the factors of fluctuations operator determinants can be calculated.

In supersymmetric gauge theories, this lengthy calculation is unnecessary. The presence of supersymmetry, with its symmetry between bosons and fermions, has the useful consequence that the fluctuations operator determinants in Eq. (6.50) in the instanton measure cancel exactly. This greatly simplifies instanton calculations in supersymmetric gauge theories, enabling exact results to be obtained in these theories purely from first principles. In Subsection 6.2.1 we shall describe the modifications which supersymmetry requires of the instanton calculus in ordinary gauge theories. Following this, in Section 6.3, we shall describe the application of the supersymmetric instanton calculus in the context of $\mathcal{N} = 2$ supersymmetric gauge theories as a means to test and fix the exact results proposed for these theories, proposed by Seiberg and Witten and subsequently generalized by others.

6.2.1 Instanton Calculus in $\mathcal{N} = 2$ Supersymmetric Gauge Theories

Instantons are classical field configurations and within the semi-classical method, valid calculations using them are restricted to weakly coupled phases. In supersymmetric gauge theories, fermions couple to bosons and the zero modes of the fermions in the instanton background must be taken into account. Furthermore, when there are scalar fields present, the situation is complicated by the possibility that these fields may possess non-zero vacuum expectation values.

Due to the presence of coupling between the gauge and matter fields and their superpartners in four dimensional supersymmetric gauge theories, the Yang–Mills instanton is no longer an exact solution of the Euler–Lagrange equations derived from the classical action of the theory. However, an approximate classical solution can be determined and then used in the semi-classical approximation. The approximate solution of the classical supersymmetric field equations is referred to as a supersymmetric instanton or super-instanton. Below we describe the $\mathcal{N} = 2$ supersymmetric ADHM instanton. Using Wick rotation, the Minkowski spacetime path integral can be analytically continued to Minkowski spacetime, enabling the semi-classical approximation and standard instanton methods to be used.

However, supersymmetric theories in four dimensional Euclidean spacetime do not admit Majorana spinors [224], and for $\mathcal{N} = 1$ supersymmetry, one must use Minkowski spacetime for semi-classical calculations (we refer also to [202, 203]). For extended supersymmetry, combinations of Weyl spinors can be formed to give Dirac spinors, which can exist in four dimensional Euclidean spacetime. Following the formalism of Chapter 3 and the review [224], we briefly describe instanton calculus for supersymmetric gauge theories in Minkowski spacetime, outlining results calculated in Euclidean spacetime and then analytically continued to Minkowski spacetime.

The presence of extended supersymmetry, or $\mathcal{N} = 1$ supersymmetry coupled to matter fields, introduces scalar fields into the field content of the theory. For scalar field theories, Derrick's theorem [78] states there can exist no non-trivial exact solutions of the classical equations of motion for theories in which the scalar field assumes a non-vanishing expectation vacuum value. This is because such solutions can always be made to vanish via a scaling argument. A scalar field possessing a non-vanishing vacuum expectation value spontaneously breaks the $SU(2)$ gauge symmetry in Seiberg–Witten theory, as described in Chapter 5. In supersymmetric gauge theories coupled to scalar fields, the super-instanton solutions of the classical equations of motion strictly do not exist, and the semi-classical approximation cannot be made. This is because the classical equations of motion involve coupled gauge fields and scalar fields. (Note that Derrick's theorem does not apply in pure gauge theories.) In such cases, Derrick's theorem can be circumvented by the 'constrained instanton' formalism formulated by Affleck [79], which was developed following a suggestion of 't Hooft. We do not describe this formalism in detail, but refer the reader to [79] and the review [224]. (Constrained instantons have been explicitly constructed in [79, 80, 213] and other references in [224]).

To leading order in the gauge coupling constant g, which is the lowest order in the semi-classical expansion, the constrained instanton is identical to the ordinary Yang–Mills instanton. When scalar vacuum expectation values are present, the boundary conditions on the scalar field are changed in the functional integral. This enables the calculation of the leading order contribution to the path integral and thus permits instanton effects in supersymmetric gauge theories to be determined.

In this chapter we have mainly followed the review [224]. We also note the fundamental work on instantons in supersymmetric theories [202, 203], the reviews [204, 205], the subsequent development of instanton calculus for supersymmetric $SU(N)$ Yang–Mills gauge field theories in [206, 207, 208], and the reviews of this in [209, 210, 211]. The modern formulation of supersymmetric instanton calculus was systematically developed in [213, 214, 217], with the general supersymmetric instanton measure first given in [222].

The $\mathcal{N} = 2$ Supersymmetric ADHM Instanton

In supersymmetric gauge theories, there exist superpartner particles, so that there are fermion fields which have consequences for instanton calculus in these theories. In $SU(N)$ gauge theories with $\mathcal{N} = 2$ supersymmetry, there are bosonic fields $\{v_m, A\}$ and fermion fields $\{\lambda_\alpha, \psi_\alpha\}$, as has already been described in Section 3.4 of Chapter 3. The fermionic superpartner of the gauge field v_m is the gaugino λ_α, which is an $\mathcal{N}=1$ superpartner field. The index theorem indicates that there are $2Nk$ zero mode solutions of the massless Dirac equation $\bar{\mathcal{D}}\lambda = 0$ in the k-instanton background. The exact form of these zero modes was derived in [41, 33], and these can be expressed in terms of the bosonic ADHM matrices $\{b, f, U\}$:

$$(\lambda_\alpha)_{uv} \;=\; \bar{U}_u^\lambda \mathcal{M}_{\lambda i} f_{ij} \bar{b}^\rho_{\alpha j} U_{\rho v} - \bar{U}_u^\lambda b_{\lambda i \alpha} f_{ij} \bar{\mathcal{M}}_j^\rho U_{\rho v}, \tag{6.51}$$

where $\mathcal{M}_{\lambda i}$ and $\bar{\mathcal{M}}_j^\rho$ are matrices of dimensions $(N+2k)\times k$ and $k\times(N+2k)$, respectively, with entries being constant Grassmann collective co-ordinates which describe the gaugino λ_α. These two matrices can be interpreted as two real Grassmann-valued matrices or as two complex Grassmann-valued matrices which are the Hermitian conjugate of each other.

Constraints can now be derived on the matrices $\mathcal{M}_{\lambda i}$ and $\bar{\mathcal{M}}_j^\rho$ by applying the Dirac operator $\bar{\mathcal{D}}$ to the gaugino in Eq. (6.51):

$$\bar{\mathcal{D}}^{\alpha\dot{\alpha}}\lambda_\alpha \;=\; 2\bar{U}b^\alpha f(\bar{\Delta}^{\dot{\alpha}}\mathcal{M} + \bar{\mathcal{M}}\Delta^{\dot{\alpha}}) f \bar{b}_\alpha U \;=\; 0. \tag{6.52}$$

Expanding the ADHM matrix $\Delta(x)$ as $\Delta(x) = a + bx$, the following fermionic constraints result from Eq. (6.52):

$$\bar{\mathcal{M}}_i^\lambda a_{\lambda j \dot{\alpha}} \;=\; -\bar{a}^\lambda_{i\dot{\alpha}} \mathcal{M}_{\lambda j}, \tag{6.53}$$

$$\bar{\mathcal{M}}_i^\lambda b^\alpha_{\lambda j} \;=\; \bar{b}_i^{\alpha\lambda} \mathcal{M}_{\lambda j}. \tag{6.54}$$

These constraints provide 'fermionic superpartners' to the bosonic ADHM constraints. The latter constraint Eq. (6.54) is straightforwardly solved by taking b to be the canonical form given in Eq. (2.63) of Subsection 2.3.1 in Chapter 2. Furthermore, the ADHM index decomposition can be applied to the fermionic matrices $\mathcal{M}_{\lambda i}$ and $\bar{\mathcal{M}}_j^\rho$. One can verify that there are $2Nk$ zero modes in these matrices. There are a total of $2k(N+2k)$ real Grassmann parameters in $\mathcal{M}_{\lambda i}$ and $\bar{\mathcal{M}}_j^\rho$; the constraints Eq. (6.53,6.53) remove $2k^2$ parameters, leaving $2Nk$ free parameters, as required.

The other superpartner fields in the theory can be treated in an analogous manner. The superpartner of the Higgs boson A, denoted ψ_α, which is also referred to as the Higgsino, can also be constructed from the bosonic and fermionic ADHM matrices, and has the form:

$$(\psi_\alpha)_{uv} \;=\; \bar{U}_u^\lambda \mathcal{N}_{\lambda i} f_{ij} \bar{b}^\rho_{\alpha j} U_{\rho v} - \bar{U}_u^\lambda b_{\lambda i \alpha} f_{ij} \bar{\mathcal{N}}_j^\rho U_{\rho v}, \tag{6.55}$$

where the matrices $\mathcal{N}_{\lambda i}$ and $\bar{\lambda}^{\rho}_{j}$ are Grassmann-valued matrices of dimensions $(N+2k)\times k$ and $k \times (N+2k)$, respectively. These matrices obey fermionic constraints analogous to Eqs. (6.53,6.53):

$$\bar{\mathcal{N}}^{\lambda}_{i} a_{\lambda j\dot{\alpha}} = -\bar{a}^{\lambda}_{i\dot{\alpha}}\mathcal{N}_{\lambda j}, \tag{6.56}$$

$$\bar{\mathcal{N}}^{\lambda}_{i} b^{\alpha}_{\lambda j} = \bar{b}^{\alpha\lambda}_{i}\mathcal{N}_{\lambda j}, \tag{6.57}$$

and the ADHM index decomposition can also be applied to these matrices.

To determine the form of the Higgs boson field A in the k-instanton background requires the solution of the Euler–Lagrange equation for this field. The field A is a complex scalar field which depends on the other superpartner fields λ and ψ, in the instanton background:

$$D^2 A = \sqrt{2}i[\lambda, \psi], \tag{6.58}$$

where D^2 is the covariant Klein–Gordon operator in the ADHM instanton background. The gaugino λ and Higgsino ψ are given by Eqs. (6.51,6.55) above. In the Coulomb phase, which is the phase of interest for gauge theories in this thesis, there is also the following long distance boundary condition on A:

$$\lim_{|x|\to\infty} A(x) = \mathrm{diag}(v_1, \ldots, v_N), \quad \sum_{u=1}^{N} v_u = 0, \tag{6.59}$$

where v_u are the complex scalar vacuum expectation values. In the $U(N)$ gauge theory, the sum of the vacuum expectation values v_u need not be zero; this is a requirement in the $SU(N)$ theory. The general solution to the equations of motion in Eq. (6.58,6.59) was first derived in [213] . The solution is complicated and is expressed in terms of the bosonic and fermionic ADHM matrices described above; A has the form:

$$iA = \frac{1}{2\sqrt{2}}\bar{U}(\mathcal{N}f\mathcal{M} - \mathcal{M}f\mathcal{N}) + \bar{U}\mathcal{A}U, \tag{6.60}$$

where $\mathcal{A}$ is a constant block diagonal matrix of dimension $(N+2k)\times(N+2k)$:

$$\mathcal{A}^{\mu}_{\lambda} = \mathcal{A}^{v+m\beta}_{u+l\alpha} = \begin{pmatrix} \langle\mathcal{A}\rangle_{uv} & 0 \\ 0 & (\mathcal{A}_{\mathrm{tot}})_{lm}\delta^{\beta}_{\alpha}, \end{pmatrix} \tag{6.61}$$

in which $\langle\mathcal{A}\rangle_{uv}$ is an $N\times N$ matrix related to the diagonal matrix of scalar vacuum expectation values:

$$\langle\mathcal{A}\rangle = i\mathrm{diag}(v_1, \ldots, v_N), \tag{6.62}$$

and $\mathcal{A}_{\rm tot}$ is an anti-Hermitian $k \times k$ matrix implicitly defined as the solution to the inhomogeneous linear equation:

$$(\mathbf{L} \cdot \mathcal{A}_{\rm tot}) \;=\; \bar{w}_{iu}^{\dot{\alpha}} \langle \mathcal{A} \rangle_{uv} w_{vj\dot{\alpha}} + \frac{1}{2\sqrt{2}} (\bar{\mathcal{M}} \mathcal{N} - \bar{\mathcal{N}} \mathcal{M})_{ij}. \tag{6.63}$$

In Eq. (6.63), $\mathbf{L}$ is a linear operator composed of ADHM submatrices which is an automorphic map on the space of $k \times k$ scalar anti-Hermitian matrices, and whose operation is given in Eq. (6.44) of Section 6.2.

Together, the constraints in Eqs. (6.53–6.63), and the ADHM constraints in Eq. (2.72,2.73) in Subsection 2.3.1 of Chapter 2, can be regarded as an $\mathcal{N} = 2$ supersymmetric multiplet of constraints which act as the fermionic superpartners of the bosonic ADHM constraints [214, 217]. For this reason these constraints are often referred to as the 'fermionic ADHM constraints.'

The field configurations defined by the bosonic and fermionic ADHM constraints serve to define the supersymmetric instanton background in the $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory. This set of configurations is referred to as the $\mathcal{N} = 2$ supersymmetric instanton, or super-instanton. The field configuration used in the semi-classical approximation in this case will consist of the leading order solutions to the equations of motion arising from the action of $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory given in Eq. (3.86) in Section 3.4 of Chapter 3.

When there are N_f fundamental matter multiplets coupled to the $SU(N)$ Yang–Mills gauge theory, the effects of the masses of these multiplets can be incorporated into the instanton effective action. The equations defining the N-dimensional fundamental matter chiral and anti-fundamental chiral superfields Q_u and $\tilde{Q}_u$ are similar to those for the other fermion fields present in the instanton background. The component fields for the chiral matter superfield Q_u are the 'Higgs' field q_u and its superpartner, the 'Higgsino,' χ_u. The fundamental fermion zero modes in the ADHM instanton background obey the following constraints:

$$\bar{\mathcal{D}} \chi_\alpha \;=\; 0, \tag{6.64}$$

$$D^2 q \;=\; -i\sqrt{2}\lambda\chi, \tag{6.65}$$

$$\lim_{|x|\to\infty} q_u(x) \;\to\; \langle q \rangle_u, \tag{6.66}$$

where $\langle q\rangle_u$ is the fundamental fermion vacuum expectation value. These constraints have solutions given by:

$$\chi_u^\alpha = \bar{U}_{u\lambda} b^\alpha_{\lambda i} f_{ij} \mathcal{K}_j, \tag{6.67}$$

$$q_u = \bar{V}_{uv}\langle q\rangle_v + \frac{i}{2\sqrt{2}} \bar{U}_{u\lambda} \mathcal{M}_{\lambda i} f_{ij} \mathcal{K}_j, \tag{6.68}$$

where $\mathcal{K}_j$ is a Grassmann number and V is the $N \times N$ submatrix of the ADHM matrix U, given in Eq. (2.90) of Subection 2.3.1 in Chapter 2.

In general, for $2N_f$ fundamental and anti-fundamental chiral matter multiplets Q_f and $\tilde{Q}_f$, one has f copies of the fields in Eq. (6.67,6.68) in the ADHM background, necessitating a further index f in these solutions. In the Coulomb phase of the theory, the fundamental fermion vacuum expectation values $\langle q\rangle_{uf}$ vanish and so Eq. (6.68) simplifies. A further modification occurs in the presence of $N_f > 0$ fundamental matter multiplets: the Euler–Lagrange equation for the conjugate Higgs field $A^\dagger$ becomes inhomogeneous, and has the form:

$$(\mathcal{D}^2 A^\dagger)_{uv} = \frac{1}{\sqrt{2}} \sum_{f=1}^{N_f} \chi_{uf} \tilde{\chi}_{fv}. \tag{6.69}$$

This equation of motion can be solved in a similar way to that for the Higgs field A, for which the Euler–Lagrange equation is unmodified. The solution, and the presence of non-zero matter multiplet masses m_f, modifies the k-instanton effective action for the $SU(N)$ Yang–Mills gauge theory, resulting in the k-instanton effective action for $\mathcal{N} = 2$ $SU(N)$ SQCD with N_f fundamental matter multiplets.

The fermion fields above which are present in the instanton background possess fermionic zero modes and also components orthogonal to these zero modes. This is directly analogous to the expansion of the fluctuations of the gauge field. The functional integral over the fermion fields can be factorized into integrations over Grassmann collective coordinates and non-zero mode components. The effective instanton action $\tilde{S}$ of the theory can then be written as a sum over the instanton action $e^{2\pi i k\tau}$, kinetic terms for the non-zero mode components S_{kin}, a term describing the interaction of the zero and non-zero modes S_{int}, and a ghost term S_{gh}. Together with the measure for the non-zero modes (which includes the non-zero mode components of the gauge field and the fermion fields) and that for the ghost fields $\{\eta, \bar{\eta}\}$, an effective action for instantons in the theory can be defined as:

$$e^{-\tilde{S}} = e^{2\pi i k\tau}\int[d\tilde{A}][d\eta][d\tilde{\eta}][d\tilde{\lambda}][d\bar{\lambda}][d\phi]e^{-S_{\rm kin}-S_{\rm int}-S_{\rm gh}}. \tag{6.70}$$

In this case, the instanton effective action is a supersymmetric potential on the instanton moduli space of collective co-ordinates.

The Collective Co-ordinate Integral in $\mathcal{N}=2$ Supersymmetric Gauge Theories

The collective co-ordinate integral requires renormalization if it is to be physically valid. A suitable renormalization scheme is the Pauli–Villars scheme [58], which is characterized by a mass scale μ. In this scheme, the cancellation between the bosonic and fermionic determinants in supersymmetric gauge theories is directly related to the scale μ:

$$\frac{\det'\Delta^{(+)}}{\det\Delta^{(-)}} = \mu^{-4Nk}. \tag{6.71}$$

The leading-order semi-classical approximation of the functional integral in the k-instanton sector is given by the supersymmetrized volume form on the instanton moduli space multiplied by an integrand involving the instanton effective action $\tilde{S}$:

$$\int[d\tilde{A}][d\eta][d\tilde{\eta}][d\tilde{\lambda}][d\bar{\lambda}][d\phi]e^{-S} \to \left(\frac{\mu}{g}\right)^{kN(4-\mathcal{N})} e^{2\pi i k\tau}\int_{\mathfrak{M}_k}\omega^{(\mathcal{N})}e^{-\tilde{S}}. \tag{6.72}$$

Here $\omega^{(\mathcal{N})}$ is the supersymmetric volume form on the instanton moduli space, which is integrated over the $U(N)$ k-instanton moduli space, denoted, as in Chapter 2, by $\mathfrak{M}_k$. This quantity has the explicit general form:

$$\begin{aligned}\int_{\mathfrak{M}_k}\omega^{(\mathcal{N})} = {} & \frac{C_k^{(\mathcal{N})}}{\mathrm{Vol}\,U(k)}\int d^{4k(N+k)}a\prod_{A=1}^{\mathcal{N}}d^{2k(N+k)}\mathcal{M}^A\,\left|\det_{k^2}\mathbf{L}\right|^{1-\mathcal{N}} \\ & \times\prod_{r=1}^{k^2}\left\{\prod_{c=1}^{3}\delta\left(\tfrac{1}{2}\mathrm{tr}_k\,T^r(\tau_{\dot\beta}^{c\dot\alpha}\bar{a}^{\dot\beta}a_{\dot\alpha})\right)\prod_{A=1}^{\mathcal{N}}\prod_{\dot\alpha=1}^{2}\delta\left(\mathrm{tr}_k\,T^r(\bar{\mathcal{M}}^A a_{\dot\alpha}+\bar{a}\dot{\alpha}\mathcal{M}^A)\right)\right\},\end{aligned} \tag{6.73}$$

where $C_k^{(\mathcal{N})}$ is a normalization factor given by:

$$C_k^{(\mathcal{N})} = 2^{-k(k-1)/2+kN(2-\mathcal{N})}\pi^{2kN(1-\mathcal{N})}. \tag{6.74}$$

As in the case of non-supersymmetric gauge theories, the supersymmetric instanton measure will exhibit the property of clustering. The supersymmetric volume form on the

$U(N)$ k-instanton moduli space $\mathfrak{M}_k$ will decompose in a directly analogous way to the non-supersymmetric volume form given in Eqs. (6.48,6.49) of Section 6.2.

Supersymmetry invariance of this volume form can be ascertained by varying the quantities within in it according to the supersymmetry transformations for the bosonic and fermionic collective co-ordinates. For extended supersymmetry, when $\mathcal{N} > 1$, the variation of the Grassmann collective co-ordinates of the fermion fields depends on the bosonic collective co-ordinates in a complicated way. The supersymmetry invariance can be established by considering the Jacobians of the supersymmetry transformations on the collective co-ordinates [222]. It has also been shown that the $\mathcal{N} = 4$ supersymmetric collective co-ordinate measure decouples to the $\mathcal{N} = 0$ supersymmetric collective co-ordinate measures via renormalization group flow [222].

The $\mathcal{N} = 2$ Supersymmetric Instanton Partition Function

We now describe the instanton partition function in $\mathcal{N} = 2$ supersymmetric gauge theories. This quantity, which for general $\mathcal{N}$-extended supersymmetry is denoted by $\mathcal{Z}_k^{(\mathcal{N})}$, is essentially a linearized reformulation of the collective co-ordinate integral. It is useful for the application of instanton calculus in these theories, and particularly for quantities which are used for comparison with proposed exact results.

One can introduce Lagrange multipliers through which both the bosonic and fermionic ADHM can be implemented as δ-function constraints; other additional variables in theories with extended supersymmetry ($\mathcal{N} > 1$) can also be included using this procedure [213, 214, 217, 222]. For the supersymmetric instanton partition function, one must introduce the auxiliary variables $\{\chi_a, \vec{D}, \bar{\psi}_A^{\dot{\alpha}}\}$. These variables consist of: a $2(\mathcal{N}-1)$-row vector of Hermitian $k \times k$ matrices, χ_a; a 3-vector of $k \times k$ Hermitian matrices, $\vec{D}$; and a $k \times k$ matrix of Grassmann-valued superpartners, $\bar{\psi}_A^{\dot{\alpha}}$, $A = 1, \ldots, \mathcal{N}$. The supersymmetric instanton partition function can be written in the form:

$$\mathcal{Z}_k^{(\mathcal{N})} = \frac{2^{2(2-\mathcal{N})k^2}\pi^{(2-3\mathcal{N})k^2}C_k^{(\mathcal{N})}}{\mathrm{Vol}\,U(k)} \int d^{4k(N+k)}a\; d^{3k^2}D\, d^{2(\mathcal{N}-1)k^2}\chi \times \prod_{A=1}^{\mathcal{N}} d^{2k(N+k)}\mathcal{M}^A\, d^{2k^2}\bar{\psi}_A\, e^{-\tilde{S}}, \tag{6.75}$$

in which the $\mathcal{N}$-supersymmetric instanton effective action is given by:

$$\tilde{S} = 4\pi^2 \mathrm{tr}_k \left\{ \chi_a \mathbf{L} \chi_a + \tfrac{1}{2} \bar{\Sigma}_{aAB} \mathcal{M}^A \mathcal{M}^B \chi_a \right\} + \tilde{S}_{\mathrm{L.m.}}, \tag{6.76}$$

and where we have defined:

$$\tilde{S}_{\mathrm{L.m.}} = -4i\pi^2 \mathrm{tr}_k \left\{ \bar{\psi}^{\dot{\alpha}}_A \left(\bar{\mathcal{M}}^A a_{\dot{\alpha}} + \bar{a}_{\dot{\alpha}} \mathcal{M}^A \right) + \vec{D} \cdot \tau^{c\dot{\alpha}}_{\dot{\beta}} \bar{a}^{\dot{\beta}} a_{\dot{\alpha}} \right\}. \tag{6.77}$$

The original form of the collective co-ordinate integral can be recovered by integrating out the auxiliary variables $\{\chi_a, \vec{D}, \bar{\psi}^{\dot{\alpha}}_A\}$ above: the auxiliary variables $\vec{D}$ and $\bar{\psi}_A$ act as Lagrange multipliers (hence the subscript "L.m."), which when integrated out give the δ-functions which implement the bosonic and fermionic ADHM constraints in Eq. (6.73). The supersymmetric instanton partition function enables one to include the effects of the vacuum expectation values of the scalar fields in the Coulomb branches of $\mathcal{N} = 2$ and $\mathcal{N} = 4$ supersymmetric gauge theories.

A particularly useful form of the supersymmetric instanton partition function is the "centred instanton partition function". This is defined in terms of an integral over the centred instanton moduli space $\mathfrak{M}_k$, defined in Eq. (2.30) of Subsection 2.2.2 in Chapter 2, in which the overall position co-ordinates and their fermionic counterparts (the superpartners) are factored off. The superpartners of the instanton are the Grassmann collective co-ordinates for the supersymmetries broken by the bosonic $U(N)$ ADHM k-instanton solution:

$$X_n = -k^{-1} \mathrm{tr}_k \, a'_n \,, \qquad \xi^A = \tfrac{i}{4} k^{-1} \mathrm{tr}_k, \mathcal{M}'^A, \tag{6.78}$$

The supersymmetric instanton effective action is always independent of X_n and ξ^A. The centered $\mathcal{N}$-supersymmetric instanton partition function $\widehat{\mathcal{Z}}^{(\mathcal{N},N_f)}_k$, generalized to include N_f matter multiplets, is given by:

$$\widehat{\mathcal{Z}}^{(\mathcal{N},N_f)}_k = \int_{\widehat{\mathfrak{M}}_k} \omega^{(\mathcal{N},N_f)} e^{-\tilde{S}}, \tag{6.79}$$

The $\mathcal{N}$-supersymmetric instanton partition function has the form of a partition function of a zero dimensional field theory. For $\mathcal{N} > 1$ it can be viewed as the dimensional reduction of the partition function of a higher-dimensional field theory.

6.3 Instanton Tests of Seiberg–Witten Theory

In this section we state the results of the tests of Seiberg–Witten theory by comparison with instanton predictions derived from first principles. We will briefly describe the results of the one-instanton and two-instanton level tests which have been performed for the exact results proposed by Seiberg and Witten. These tests make extensive use of the supersymmetric instanton calculus described in Section 6.2. We do not describe the techniques employed to determine the instanton contributions to the prepotential or Green's functions in Seiberg–Witten theory and its generalizations in detail but shall state results. The procedure conventionally employed in testing the proposed exact results by comparing them with instanton predictions is to calculate the instanton corrections to Green's functions using the semi-classical approximation. One of these results shall then be used, in Section 6.4, for the purposes of matching the one-instanton prediction and the proposed exact result for the prepotential in low energy effective $\mathcal{N} = 2$ $SU(N)$ SQCD with $N_f = 2N$ massless fundamental matter multiplets.

Instanton tests of the proposed generalization of Seiberg–Witten theoretic methods to determine the exact low energy effective action of $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with N_f fundamental matter multiplets shall be described in Section 6.4.

In the following titled paragraphs we will describe the general procedure for the calculations whose results we state for Seiberg–Witten theory and for $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories in Section 6.4 below. We describe the results of the one-instanton and two-instanton tests of the original Seiberg–Witten theory, namely $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, and $\mathcal{N} = 2$ $SU(2)$ SQCD with $N_f \leq 4$ fundamental matter multiplets. In Section 6.4 we describe the results of the one-instanton tests of the exact results proposed for the theory of primary focus in this thesis, $\mathcal{N} = 2$ $SU(N)$ SQCD with $N_f \leq 2N$ fundamental matter multiplets.

We begin by first describing the methods used to calculate the instanton contributions to the prepotential $\mathcal{F}$ in $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory with $N_f \leq 4$ fundamental matter multiplets, the object for which Seiberg and Witten have proposed an exact low energy effective form. In this section we refer to the original papers in [212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 223, 225, 226, 227, 228], and note the related work in [241, 242, 243, 244, 247] on instanton tests of Seiberg–Witten theory and its generalizations.

The initial instanton tests of Seiberg–Witten theory were performed for the one-instanton contributions to the prepotential in $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, the first one being that in [212]. This was followed by one-instanton tests for the exact solution of $\mathcal{N} = 2$ $SU(2)$ SQCD with N_f fundamental matter multiplets proposed by Seiberg and Witten [221, 218].

The first test of Seiberg–Witten theory at the two-instanton level was completed by Dorey et al. [213]. This calculation was generalized to the case of $\mathcal{N} = 2$ $SU(2)$ SQCD coupled to N_f fundamental matter multiplets in [214, 217, 215]. We state the results of the calculations completed for the one-instanton and two-instanton contributions to the prepotential following [224]. The one-instanton results are in agreement with the predictions from Seiberg–Witten theory, for $N_f \leq 2N - 1$, except for the finite scale invariant case when $N_f = 2N$. This discrepancy and its proposed resolution shall be described in Section 6.5. At the two-instanton level, the tests performed so far have also agreed with Seiberg–Witten theory [214, 217].

The general expression for the k-instanton contribution to the prepotential in $\mathcal{N} = 2$ $SU(2)$ SQCD as an integral over the instanton moduli space was also derived in [214, 217]. This constitutes a complete field theoretic solution for the low energy Wilsonian effective of $\mathcal{N} = 2$ $SU(2)$ SQCD with N_f matter multiplets given in terms of quadratures.

The Matone relation [229], described in Chapter 5, which relates the k-instanton contributions to the prepotential to the quantum modulus, or condensate, u_2, in Seiberg–Witten theory, has also been subject to tests by independent instanton calculations. These tests also provide checks of the Seiberg–Witten proposed solution via instanton calculus. The Matone relation was tested at the two-instanton level in [230]. This was later extended to an all orders k-instanton test in [231].

Instanton Contributions to the Prepotential

In the Coulomb phase of $\mathcal{N} = 2$ supersymmetric gauge theories, there exists a scalar field

ϕ transforming in the adjoint representation of the gauge group whose non-zero vacuum expectation values break the full gauge symmetry of the theory. Generically, the gauge group is broken to its maximal Abelian subgroup. When the scalar vacuum expectation value, which is the only vacuum expectation value on the Couloumb branch, is large, the gauge theory becomes weakly coupled. At weak coupling, instanton calculations are expected to be applicable and reliable. The instanton calculus in supersymmetric gauge theories, described in Subsection 6.2.1, making use of constrained instantons, can then be applied to these theories in the Coulomb phase.

The presence of $\mathcal{N} = 2$ supersymmetry protects holomorphic quantities in these theories from quantum perturbative corrections beyond one loop corrections. As has been described in Chapter 5, the prepotential of the low energy effective $\mathcal{N} = 2$ theory is one such holomorphic quantity. This protection from all higher orders of perturbative contributions then enables the exact identification of non-perturbative contributions, which are otherwise often negligible in comparison to the infinite number of perturbative contributions. Hence $\mathcal{N} = 2$ supersymmetric gauge theories are particularly suitable for the use of the semi-classical methods, which are expected to be exact in these theories.

The instanton contributions to the prepotential of the theory are conventionally calculated using the semi-classical approximation for certain Green's functions. The leading order contributions to these Green's functions are calculated and then related to the low energy effective action. The Green's functions are such that in the long distance (infra-red) limit they can be compared with the appropriate expansion of the low energy effective action proposed by Seiberg and Witten or a generalization of it. This enables one to test the prepotential proposed for the low energy effective action. Various Green's functions can be used for such tests, such as the four-point anti-chiral fermion correlator which depends on the fourth derivative of the prepotential with respect to the scalar vacuum expectation value. Other correlators exist which can be used to determine the second derivative of the prepotential.

When calculating the instanton contributions to the $\mathcal{N} = 2$ prepotential and quantum moduli for the purposes of testing the proposed exact results in Seiberg–Witten theory and its generalizations, the semi-classical approximation is used to calculate quantities in $\mathcal{N} = 2$ SQCD with N_f fundamental matter mutliplets. With results for this theory, for various gauge groups, testing the $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory is achieved by setting the number of matter multiplets to zero, $N_f = 0$. This is the theory considered in the majority of literature on this topic, and we follow this approach here.

We now outline the form of the $\mathcal{N} = 2$ supersymmetric instanton effective action $\tilde{S}$ for $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge theory, with general gauge group, in the Coulomb phase. The scalar field ϕ obeys the same equation of motion as in the pure Yang–Mills theory, and these can be used to simplify the kinetic terms for the scalar fields which appear in the instanton effective action. The kinetic terms of the scalar fields and the Yukawa interactions for the chiral fermions contribute to leading order in $\tilde{S}$. On the Coulomb branch of the theory the matter multiplet fermion fields do not acquire vacuum expectation values, that is, they have vanishing vacuum expectation values. The only field which possesses a vacuum expecation value is the scalar field ϕ. This also simplifies the form of $\tilde{S}$, leaving the terms involving the scalar field ϕ as the only non-vanishing terms in the integrand. This permits the contribution of ϕ to $\tilde{S}$ to be calculated.

The Yukawa interaction terms can also be evaluated by similar techniques. Summing the contributions of the scalar fields and the Yukawa interaction terms, the leading order expression for the instanton effective action $\tilde{S}$ in $\mathcal{N} = 2$ Yang–Mills gauge theory on the Coulomb branch (with non-zero scalar field vacuum expectation values) results. The effective action $\tilde{S}$ is a supersymmetric invariant, as required.

The addition of $\mathcal{N} = 2$ fundamental matter multiplets to the pure $\mathcal{N} = 2$ gauge theory affects the instanton effective action of the theory. To leading order in the semi-classical approximation, this can be achieved relatively straightforwardly.

The quantities of interest in these theories, which includes the $\mathcal{N} = 2$ prepotential $\mathcal{F}$, are holomorphic in the matter multiplet masses. That is, they depend on a mass m but not its conjugate m^*, where m is a complex mass. The variables m and m^* can as independent variables so that $m^* = 0$ can be set. When calculating these quantities using the semi-classical approximation, the holomorphic mass dependence is manifest in the approximation. Due to the holomorphic mass dependence, only the equations of motion for anti-chiral fermions in the theory are affected. Hence the equations defining the supersymmetric instanton are not affected at leading order.

The effect of including fundamental matter multiplets in the theory is accounted for as follows. The mass term in the action involving a fundamental hypermultiplet transforming in the $(N, \bar{N})$ representation can be evaluated in the background of the supersymmetric instanton configuration. This is then added to the effective instanton action $\tilde{S}$, and in this way the mass terms for the matter multiplets has been accounted for. A similar procedure permits one to account for the effect of coupling the $\mathcal{N} = 2$ Yang–Mills gauge to an adjoint matter multiplet. This is useful for calculating instanton contributions to the mass-deformed $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory described in Chapter 5 and in Section 6.5.

Using the $\mathcal{N} = 2$ supersymmetric effective instanton action $\tilde{S}$, one can semi-classically calculate the four-point anti-chiral fermion correlator at long range. When this result is compared with the same quantity predicted by the proposed exact results, the k-instanton expansion coefficient of the prepotential assumes the form:

$$\mathcal{F}_k = g^{-k(2N-N_f)+2} \widehat{\mathcal{Z}}_k^{(\mathcal{N}=2,N_f)}, \tag{6.80}$$

where $\widehat{\mathcal{Z}}_k^{(\mathcal{N}=2,N_f)}$ is the $\mathcal{N} = 2$ supersymmetric centered instanton partition function. This relation is valid up to an undetermined additive constant, which does not affect the physics described by the prepotential $\mathcal{F}$. This is because only derivatives of the prepotential appear in the low energy effective action of the theory. Furthermore, since only fourth derivatives of $\mathcal{F}$ enter into the correlator from which this relation is derived, other functions whose fourth derivative with respect to the scalar vacuum expectation values is a constant which can be added to it. However, relations derived from other correlators permit only functions whose second derivative with respect to the scalar vacuum expectation values vanishes to be added.

Instanton Contributions to u_2

An alternative approach for testing Seiberg–Witten theory against instanton predictions does not use Green's functions. Instead, the instanton contributions to the quantum modulus, or condensate, u_2, can be calculated and related to the prepotential by using a renormalization group equation involving the derivative of the prepotential. This calculation also serves to verify this renormalization group relation.

In the alternative approach which uses the renormalization group equation known as the Matone relation in Seiberg–Witten theory, the k-instanton contributions to the quantum modulus u_2 can be calculated using the semi-classical approximation as in the previous method. The authors of [224] evaluate the instanton contributions to u_2 in $\mathcal{N}=2$ $SU(N)$ SQCD with N_f fundamental matter multiplets. The k-instanton contribution to the quantum modulus u_2 in this case is given by:

$$u_2|_k \;=\; 2\pi k \Lambda_{(N_f)}^{k(2N-N_f)} g^{-k(2N-N_f)+2}\, \widehat{\mathcal{Z}}_k^{(\mathcal{N}=2,N_f)}. \tag{6.81}$$

The k-instanton contribution to the quantum modulus u_2 is proportional to the centred instanton partition function, as is the k-instanton contribution to the prepotential $\mathcal{F}$ in Eq. (6.80). This suggests a version of the renormalization group equation analogous to the Matone relation, but for the gauge group $SU(N)$. By comparing Eq. (6.81) with Eq. (6.80), one can relate the quantum modulus u_2 to the prepotential $\mathcal{F}$ as follows:

$$u_2|_k \;=\; 2\pi k \Lambda^{k(2N-N_f)} \mathcal{F}_k, \tag{6.82}$$

which generalizes the original Matone relation given previously in Eq. (5.201) of Subsection 5.3.2 in Chapter 5. This formula relates only the non-perturbative contributions to both u_2 and $\mathcal{F}$, so that there remain perturbative contributions to be determined. The prepotential $\mathcal{F}$ receives one loop perturbative corrections, which can be calculated using standard perturbative methods, and which are given in Eq. (5.40) in Section 5.3 of Chapter 5. By itself, the quantum modulus u_2 receives no perturbative corrections and is equal to its classical form.

This alternative approach has the considerable advantage that the relation Eq. (6.82) can be derived without integrating over the instanton moduli space, thus avoiding the ADHM constraints, via instanton calculus. This relation is also valid for the finite scale invariant case $N_f = 2N$ if one replaces the dynamical scale factor Λ^{2N-N_f} with the exponentiated factor $e^{2\pi i \tau}$. Equation (6.82) also holds for arbitrary matter multiplet masses.

We note that the instanton sectors which contribute to the prepotential are dependent on the theory being considered. Due to a $\mathbb{Z}_2$ parity symmetry in the theory with $N_f > 0$ massless fundamental matter multiplets, described in Chapter 5, it is found that only instantons of even charge contribute to the prepotential in these theories. In the case of the $\mathcal{N}=2$ supersymmetric Yang–Mills gauge theory, with $N_f=0$, all instantons, of both odd and even charge, contribute to the prepotential. This is a selection rule which was discovered by Seiberg and Witten [171], and which can also be derived from instanton calculus. When some of the $N_f>0$ fundamental matter multiplets have non-zero mass, this symmetry is absent.

One-Instanton Test

We now describe the actual instanton tests which have been performed using the approach utilizing Green's functions. Using precisely this approach for the gauge group $SU(2)$, the authors [212] were able to calculate the one-instanton contribution to the prepotential $\mathcal{F}$. These one-instanton tests constitute the first tests of Seiberg–Witten theory by conventional field theoretic methods. Other one-instanton tests have been performed for Seiberg–Witten theory, including [221, 218]. Further one-instanton tests of cases involving $\mathcal{N}=2$ supersymmetric $SU(2)$ Yang–Mills gauge theory coupled to assorted matter multiplets have also been completed [227].

The formalism used for instanton calculus so far has been that for the gauge group $U(N)$, which is also applicable for gauge group $SU(N)$. For the gauge group $SU(2)$, it is more convenient and economical to make use of the isomorphism $SU(2)\simeq Sp(1)$ and use the ADHM construction for the sympletic groups described in Subsection 2.3.3 of Chapter 2. The differences between the formalism of the ADHM construction for the unitary and sympletic gauge groups has already been described in Subsection 2.3.3 of Chapter 2; here we recall that the ADHM construction for gauge group $Sp(1)$ requires a smaller number of variables and constraints than the ADHM construction for gauge group $SU(2)$. The most simple choice of formalism with which to proceed for the direct calculation of the $SU(2)$ prepotential is that for $Sp(1)$. However, using the $SU(N)$ instanton calculus with $N=2$ should reproduce the same physical results and would be equally valid.

Calculation of the k-instanton contribution to the prepotential, denoted $\mathcal{F}_k$, requires the explicit centred instanton partition function. For the one-instanton contribution to the prepotential of $\mathcal{N}=2$ supersymmetric $SU(2)$ Yang–Mills gauge theory, denoted $\mathcal{F}_1|_{N=2,N_f=0}$, which is the simplest case, the bosonic and fermionic parameters of a $k=1$

$\mathcal{N} = 2$ supersymmetric ADHM instanton are specified in three 2×1 matrices of independent, unconstrained quaternionic parameters:

$$a = \begin{pmatrix} w \\ -X \end{pmatrix}, \quad \mathcal{M}^A = \begin{pmatrix} \mu^A \\ -4i\xi^A \end{pmatrix}. \tag{6.83}$$

For the $Sp(1)$ ADHM one-instanton, as noted in Subsection 2.3.3 of Chapter 2, there are no ADHM constraints, so that the ADHM matrix a in Eq. (6.83) completely specifies the $Sp(1)$ one-instanton. There are also $2N_f$ Grassmann variables $\{\mathcal{K}_f, \tilde{\mathcal{K}}_f\}$ which parameterize the zero modes of the fundamental matter multiplets.

The volume form on the centered instanton moduli space, for use in the centered instanton partition function, is given by:

$$\int_{\widehat{\mathcal{M}}_1} \omega^{(\mathcal{N}=2,N_f)} = \frac{2^3}{\pi^{4+2N_f}} \int d^4u \prod_{A=1}^{2} d^2\mu^A \prod_{f=1}^{N_f} d\mathcal{K}_f\, d\tilde{\mathcal{K}}_f, \tag{6.84}$$

The $k = 1$ instanton effective action can be determined from this and used in the calculation of the supersymmetric one-instanton partition function, $\widehat{\mathfrak{Z}}_1^{(\mathcal{N}=2,N_f)}$. The result for this quantity is:

$$\widehat{\mathfrak{Z}}_1^{(\mathcal{N}=2,N_f)}|_{N=2} = \frac{2}{a^2} \prod_{f=1}^{N_f} m_f, \tag{6.85}$$

where a is the non-zero scalar vacuum expectation value of the theory, thus implying that the theory is in the Coulomb phase. The form of the one-instanton contribution to the prepotential in $\mathcal{N} = 2$ supersymmetric $SU(2)$ SQCD with N_f fundamental matter multiplets can then be found using Eq. (6.80). For the case of $N_f = 0$, which is the arena of the original low energy solution proposed by Seiberg and Witten, it follows that the one-instanton contribution to the prepotential, $\mathcal{F}_1|_{N=2,N_f=0}$, as given in [224], has the form:

$$\mathcal{F}_1|_{N=2,N_f=0} = \frac{2}{a^2}, \quad \mathcal{F}_1|_{N=2,N_f} = \frac{2}{a^2} \prod_{f=1}^{N_f} m_f \tag{6.86}$$

The expressions for $\mathcal{F}_1|_{N=2,N_f=0}$ and $\mathcal{F}_1|_{N=2,N_f}$ in Eq. (6.86) are in exact agreement with the form of $\mathcal{F}_1|_{N=2,N_f}$ derived from Eq. (5.115) of Chapter 6 for $0 \leq N_f < 4$, up to constants independent of the vacuum expectation value a for the cases $N_f = 2, 3$. When $N_f = 4$, a discrepancy arises between the results of the instanton calculation and the

Seiberg–Witten result for $\mathcal{F}_1|_{N=2,N_f=0}$. We shall describe the resolution of this discrepancy, which is achieved by a non-perturbative reparameterization, in Section 6.5.

Two-Instanton Test

The calculation of the two-instanton contribution to the prepotential, which for $\mathcal{N}=2$ supersymmetric $SU(2)$ Yang–Mills gauge theory we denote by $\mathcal{F}_2|_{N=2,N_f=0}$, presents more difficulties than the one-instanton contribution. The bosonic and fermionic parameters of the $k=2$ supersymmetric ADHM instanton are contained in the following 3×2 quaternionic matrices:

$$a = \begin{pmatrix} w_1 & w_2 \\ r_{11} & r_{12} \\ r_{12} & r_{12} \end{pmatrix}, \quad \mathcal{M}^A = \begin{pmatrix} \mu_1^A & \mu_2^A \\ -4i\xi^A + \mathcal{M}_3'^A & \mathcal{M}_1'^A \\ \mathcal{M}_1'^A & -4i\xi^A - \mathcal{M}_3'^A \end{pmatrix}, \tag{6.87}$$

the bosonic components of which, contained in the ADHM matrix a, have been described in Subsection 2.3.3 of Chapter 2. Note that for the $Sp(1)$ ADHM matrix a in Eq. (6.87), a redefinition of the diagonal parameters, as originally defined in Subsection 2.3.3 of Chapter 2, has been made. The elements of the matrix $\mathcal{M}^A$ are Weyl spinors. There are also $4N_f$ fundamental zero modes parameterized by the Grassmann numbers $\mathcal{K}_{if}$ and $\tilde{\mathcal{K}}_{fi}$, which is twice the number of such modes in the one-instanton calculation.

Again the the centred instanton partition function is to be evaluated in order to determine the instanton contributions. The first step in this calculation is the explicit solution of the bosonic and fermionic ADHM constraints. This is achieved by using the $Sp(1)$ ADHM two-instanton solution given in Subsection 2.3.3 of Chapter 2, and using an analogous technique for the fermionic ADHM constraints, in which the off-diagonal element $\mathcal{M}_1'^A$ is eliminated. The constraints for the bosonic and fermionic ADHM supersymmetric $Sp(1)$ two-instanton are then explicitly and generally solved by:

$$r_{12} = \frac{1}{2|X|^2} X\,(\bar{w}_2 w_1 - \bar{w}_1 w_2), \quad \mathcal{M}_1'^A = \frac{1}{2|X|^2} X\,(2\bar{r}_{12}\mathcal{M}_3'^A + \bar{w}_2\mu_1^A - \bar{w}_1\mu_2^A), \tag{6.88}$$

where the real constant (denoted Σ in Chapter 2) which can be added to the the term $(\bar{w}_2 w_1 - \bar{w}_1 w_2)$ in r_{12} has been set to zero via the $O(2)$ residual symmetry, and we have defined $X \equiv r_{11} - r_{22}$. This configuration can now be used to explicitly integrate the

δ-functions which implement the supersymmetric ADHM constraints, in the formula for the supersymmetric centred $k = 2$-instanton volume form:

$$\int_{\widehat{\mathcal{M}}_2} \omega^{(\mathcal{N}=2)} = \frac{C_2^{(\mathcal{N}=2)}}{2^{10}} \int d^4 r'_{12}\, d^4 w_1\, d^4 w_2 \left\{ \prod_{A=1}^{2} d^2 \mathcal{M}_3'^A\, d^2 \mu_1^A\, d^2 \mu_2^A \right\} \times \frac{\left| |X|^2 - |r_{12}|^2 \right|}{|w_1|^2 + |w_2|^2 + 2|r_{12}|^2 + 2|X|^2}. \tag{6.89}$$

The bosonic components of the general $Sp(1)$ two-instanton collective co-ordinate integral were first derived in [47, 44, 46] by an explicit change of variables in the path integral. The resulting instanton measure is known as the Osborn measure [47]. The explicit form of the collective co-ordinate measure for higher instanton numbers is not known.

After a lengthy calculation using the instanton calculus which involves many subtleties, the two-instanton contribution to the prepotential in $\mathcal{N} = 2$ $SU(2)$ SQCD with N_f fundamental matter multiplets is given by [213, 214, 217, 224]:

$$\mathcal{F}_2|_{N=2,N_f} = \frac{5}{a^6} M_{N_f}^{(N_f)} - \frac{3}{4a^4} M_{N_f-1}^{(N_f)} \cdot \frac{1}{16a^2} M_{N_f-2}^{(N_f)} - \frac{5}{2^6 3^3} M_{N_f-3}^{(N_f)} + \frac{7a^2}{2^8 3^5} M_{N_f-4}^{(N_f)}, \tag{6.90}$$

where the coefficients $M_l^{(N_f)}$ are defined

$$M_l^{(N_f)} \equiv \sum_{f_1 < f_2 < \cdots < f_l = 1}^{N_f} m_{f_1}^2 m_{f_1}^2 \cdots m_{f_l}^2, \tag{6.91}$$

in which the factors m_{f_l} are the masses of the l^{th} fundamental matter multiplet, $l = 1, \ldots, N_f$. The sum in Eq. (6.91) is defined such that $M_0^{(N_f)} = 1$, and $M_l^{(N_f)} = 0$ for $l < 0$. For $N_f < 4$, the two-instanton contribution to the $SU(2)$ prepotential $\mathcal{F}_2|_{N=2,N_f}$ given in Eq. (6.90) is in exact agreement with the Seiberg–Witten prediction given explicitly in Eq. (5.115) in terms of the complexified coupling τ, up to a physically unimportant additive constant in the case $N_f = 3$.

6.4 Instanton Tests of the Exact Results in $\mathcal{N} = 2$ Supersymmetric $SU(N)$ Gauge Theory

The analysis which led to the results stated above for the $SU(2)$ theory can be generalized to the general case in which the theory has gauge group $SU(N)$. The $SU(N)$ instanton partition function previously defined in Eq. (6.75) of Subsection 6.2.1 can be used to calculate the instanton contributions to the prepotential via the Green's function method making use of the formula Eq. (6.80). The calculations proceed similarly but prove more lengthy and complicated. The ADHM construction used for the bosonic and fermionic components of the $\mathcal{N}=2$ supersymmetric one-instanton must now be that for the unitary groups $U(N)$ with the necessary modifications made to account for the presence of matter multiplets. The formalism for the $U(N)$ ADHM construction was described in Subsection 2.3.1 of Chapter 2, and is more complicated than the formalism for the $Sp(1) \simeq SU(2)$ ADHM construction.

We first state the result of the calculation completed for the one-instanton contribution to the prepotential in $\mathcal{N}=2$ $SU(N)$ SQCD with N_f fundamental matter multiplets. The details of this difficult calculation can be found in [223]. The underlying method used to complete the calculation is precisely the same as that employed for the gauge group $Sp(1) \simeq SU(2)$ described above. We note that other authors have investigated the instanton contributions in the $SU(N)$ theory [218].

In the Coulomb phase, $\mathcal{N}=2$ supersymmetric $SU(N)$ Yang–Mills gauge field theories possess an $N-1$ complex dimensional classical moduli space which is parameterized by the N scalar vacuum expectation values $\{a_i\}$, $i=1,\ldots,N$ of the N scalar fields ϕ_i. Generically, on the Coulomb branch, at each point of the classical moduli space the gauge group $SU(N)$ is broken to its maximal Abelian subgroup $U(1)^{N-1}$. In this phase, the long distance low energy effective $SU(N)$ action S_{eff} can be expressed in terms of Abelian $\mathcal{N}=1$ superfields $W_{\alpha i}=((v_m)_i,\lambda_i)$, $\Phi_i=(\phi_i,\psi_i)$ and the dual superfield $\Phi_{\mathrm{D}i}(\Phi)$. The form of S_{eff} is then given by:

$$S_{\text{eff}} = \frac{1}{4\pi}\int d^4x\,\mathrm{Im}\left\{\tfrac{1}{2}\tau_{ij}(\Phi)\,W_i^{\alpha}W_{\alpha j}\Big|\theta^2 + \Phi_{\mathrm{D}i}(\Phi)\;\Phi_i^{\dagger}\Big|_{\theta^2\bar{\theta}^2}\right\} \tag{6.92}$$

In terms of these $\mathcal{N}=1$ superfields, the scalar components of Φ_i are given by $g\phi_i/\sqrt{2}$. The low energy effective action S_{eff} is determined uniquely by the prepotential $\mathcal{F}(\Phi)$ [170]. The dual scalar superfield $\Phi_{\mathrm{D}i}$ and matrix of complexified gauge couplings τ_{ij} are related to the prepotential in a way which generalizes that for the $SU(2)$ theory:

$$\Phi_{\mathrm{D}i} \equiv \frac{\partial\mathcal{F}}{\partial\Phi_i},\quad \tau_{ij} \equiv \frac{\partial^2\mathcal{F}}{\partial\Phi_i\partial\Phi_j} \tag{6.93}$$

The matrix τ_{ij} of the low energy effective $U(1)^{N-1}$ theory then depends upon the scalar vacuum expectation values $\{a_i\}$. The scalar vacuum expectation values and their dual values serve as local co-ordinates in different regions of the quantum moduli space. Where $\{a_i\}$ are the local co-ordinates, these values are functions of the quantum moduli u_n, $a_i = a_i(u_n)$. Similarly, where the dual scalar fields $\{\phi_{\mathrm{D}i}\}$, which are the scalar components of the dual superfield $\Phi_{\mathrm{D}i}$, parameterize the quantum moduli space, the dual values are also functions of the quantum moduli: $a_{\mathrm{D}i} = a_{\mathrm{D}i}(u_n)$. The long distance behaviour of correlation functions can be extracted from the low energy effective $SU(N)$ action given in Eq. (6.92). The correlation function which was used to determine the one-instanton contribution to the $SU(N)$ prepotential, given below, is again the Euclidean four-point anti-chiral fermion correlator. This particular Green's function was used for the derivation of the one- and two-instanton contributions to the prepotential in the $SU(2)$ theory, as described in Section 6.3. The coefficients of the instanton contributions in the $SU(N)$ theory will also depend upon the scalar vacuum expectation values, and, when there are N_f fundamental matter multiplets present, upon the multiplet masses.

One-Instanton Test

Extensive studies of the one-instanton contributions to the prepotential of $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory coupled to N_f fundamental matter multiplets, which we denote as $\mathcal{F}_1|_{N,N_f}$, have been made in [218, 223]. We state the result derived in [223], which is valid for $\mathcal{N} = 2$ $SU(N)$ SQCD with N_f massive or massless fundamental matter multiplets.

The Green's function used to evaluate the one-instanton contribution to the $SU(N)$ prepotential makes use of the instanton partition function given in Eq. (6.75) in Subsection 6.2.1. The integrations to be performed are not elementary, contain many subtleties, and those over the $SU(N)$ one-instanton moduli space are found to be highly complex and non-trivial. Complete details of this lengthy calculation can be found in [223, 224].

The $k = 1$ contribution to the centred partition function in $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory coupled to N_f fundmental matter multiplets, each of mass m_f, $f = 1, \ldots, N_f$, is given by:

$$\mathcal{F}_1|_{N,N_f} = -\frac{iC_1'\pi^{2N-1}}{2^{N+2}} \sum_{i=1}^{N} \sum_{i \neq j} \frac{1}{(a_i - a_j)\gamma_i\gamma_j} \cdot \prod_{f=1}^{N_f} \left(-\frac{1}{\sqrt{2}}(a_i + a_j) + m_f \right), \tag{6.94}$$

where we have defined the factors:

$$\gamma_i = \prod_{k \neq i, k \neq j} (a_i - a_k). \tag{6.95}$$

This result agrees exactly with the results of the generalizations of Seiberg–Witten theory to the $SU(N)$ case in [173, 174, 181, 182, 183, 233], for $N_f < 2N$. An example of the hyperelliptic curves proposed for the $SU(N)$ theory was given in Eq. (5.214). For the scale invariant theories with $N_f = 2N$ fundamental matter multiplets, there exist discrepancies which we shall describe in Subsection 6.5.1 of Section 6.5.

Two-Instanton Test

The possibility of a two-instanton test in the $SU(N)$ theory using the previously described methods has recently been enhanced by the determination of the explicit exact general form of the $U(N)$ ADHM two-instanton configuration [36]. This work has been described in detail in Subsection 2.3.2 of Chapter 2. The complexity of the $U(N)$ ADHM two-instanton configuration implies that the calculation of the two-instanton contributions to the prepotential in $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory coupled to N_f fundamental matter multiplets, denoted by $\mathcal{F}_2|_{N,N_f}$, would be technically demanding. However, aside from this a two-instanton calculation in the $SU(N)$ theory presents additional difficulties [223]. This is because in general there exist $2N(k-1)$ relative fermionic zero modes for a k-instanton calculation in the $SU(N)$ theory. These additional zero modes do not correspond to Lagrangian symmetries and therefore are less readily identified [223]. In the one-instanton calculation, which yields $\mathcal{F}_1|_{N,N_f}$ in Eq. (6.94), these additional zero modes vanish and do not contribute.

6.5 Matching Exact Results and Instanton Predictions

In this section we describe the matching of exact field theoretic predictions from instanton calculations for supersymmetric gauge theories and their proposed exact counterparts derived using Seiberg–Witten theory or its generalizations.

As has been described in Section 6.3, instanton calculus provides a means for the calculation of the instanton contributions to the low energy effective prepotential $\mathcal{F}$ in Seiberg–Witten theory and other $\mathcal{N} = 2$ supersymmetric gauge theories for which exact results have been proposed. Specifically, instanton calculus can be used to evaluate the numerical coefficients $\mathcal{F}_k$, where k is the instanton charge, which occur in the non-perturbative contributions to $\mathcal{F}$. Other tests of the Seiberg–Witten solution via instanton calculations were also described, such as that for the instanton contributions to the quantum modulus u_2.

For certain gauge groups and numbers of matter multiplets, the distinct methods for calculating the coefficients $\mathcal{F}_k$ do not agree. Furthermore, some of the assumptions made by Seiberg and Witten in deriving their solution would appear to be invalid as non-perturbative contributions have been assumed not to exist. This is true in the when one equates the classical complexified gauge coupling constant τ_{cl} to its quantum counterpart, τ_{qu}. When matter multiplets are present, the equality $\tau_{\text{cl}} = \tau_{\text{qu}}$ does not hold, as this relation receives non-perturbative quantum corrections from instantons.

Such instances of disagreement between instanton predictions and the proposed exact results show that the exact non-perturbative quantities proposed in Seiberg–Witten theory and its generalizations must be matched to field theoretic calculations at low instanton charge if they are to describe the correct non-perturbative physics at higher instanton charge. Instanton contributions to quantities such as the complexified gauge coupling τ must also be taken into account if the proposed exact results are to describe the non-perturbative regime in the same way as field theoretic calculations do.

Below we briefly describe previous work which reports discrepancies between the proposed exact results and instanton calculations. The most serious discrepancies occur for scale invariant theories. In the theories considered by Seiberg and Witten, these theories include $\mathcal{N} = 2$ $SU(2)$ SQCD with $N_f = 4$ fundamental matter multiplets, $\mathcal{N} = 2$ $SU(2)$ Yang–Mills gauge theory with an adjoint $\mathcal{N} = 2$ matter multiplets (previously referred to mass deformed $\mathcal{N} = 4$ Yang–Mills gauge theory), and more generally, $\mathcal{N} = 2$ $SU(N)$ SQCD with $N_f = 2N$ fundamental matter multiplets. The original analysis of Seiberg and Witten for these theories with $N_f \leq 3$ cannot be applied in these cases because there is no running of the gauge coupling or the energy scale, making the determination of the elliptic curves for the quantum moduli space more difficult. Analogous difficulties occur in the generalizations of the analysis for larger gauge groups, and in particular we will shall describe the case with gauge group $SU(N)$.

The results for the low energy effective actions derived from the proposed exact solutions were found not to agree with instanton predictions in these theories for $N_f = 2N$ [217, 218, 219, 223]. There are also discrepancies in the expressions for the quantum moduli for $N < N_f < 2N$, as reported in [215, 218, 221, 223].

In Subsection 6.5.1, a one-instanton level test is performed for the proposed reparameterization scheme matching the conjectured exact low energy results and instanton predictions for $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories with $2N$ massless fundamental matter hypermultiplets across the entire quantum moduli space. The constants within the scheme which ensure agreement between the exact results and the instanton predictions for general N are derived. This constitutes a non-trivial test of the scheme, which eliminates the discrepancies arising when the two sets of results are compared.

Using the results derived via instanton calculations stated in Section 6.3, precise agreement between the predictions of Seiberg–Witten theory have been obtained for $\mathcal{N} = 2$ $SU(2)$ SQCD with $N_f \leq 3$ fundamental matter multiplets. Disagreement has been claimed for the case $N_f = 3$, but this has been resolved [216].

$\mathcal{N} = 2$ $SU(2)$ SQCD with $N_f = 4$ Matter Multiplets

There are three distinct instances of disagreement between the exact results proposed by Seiberg and Witten for $\mathcal{N} = 2$ $SU(2)$ SQCD with $N_f = 4$ fundamental matter multiplets and instanton calculations. We shall describe each of these in turn. No modification of the elliptic curve proposed to describe the quantum moduli space of vacua of the theory, given in Eq. (5.188) of Chapter 5, is required. However, a finite non-perturbative renormalization of the parameters τ and u in the theory is necessary. This alters the relation between the effective values of these parameters and their values in the microscopic (high energy) theory. In short, the quantities which Seiberg and Witten take in their solution to be equal to their classical values for this finite theory are actually subject to quantum non-perturbative corrections.

In the case when there are N_f fundamental matter multiplets in the $\mathcal{N} = 2$ $SU(2)$ SQCD theory, those models with $N_f \leq 4$, one multiplet of which is massless, are enhanced by an $\mathbb{Z}_2$ parity symmetry. This symmetry forbids the contribution of odd instanton charge configurations to the prepotential of these theories, as noted in the Section 6.3.

Firstly, Seiberg and Witten assume that in the scale invariant theory with $N_f = 4$, in which there is no running of the gauge coupling constant, the effective coupling τ_{eff} is equal to the classical gauge coupling τ of the full $SU(2)$ gauge theory:

$$\tau_{\text{eff}} \;=\; \tau. \tag{6.96}$$

However, Dorey et. al [219] determine the correct form of the effective complexified gauge coupling constant τ_{eff} of the $U(1)$ gauge theory in Seiberg–Witten theory from first principles using instanton calculus. In fact, as is shown in [219], the Seiberg–Witten elliptic curve for this theory, given in Eq. (5.188), is parameterized by an effective complexified gauge coupling τ_{eff} instead of the microscopic coupling τ. The result is that the effective gauge coupling τ_{eff} receives an infinite series of non-perturbative corrections from instantons of even charge, by which it differs from the microscopic coupling τ. The corrected effective gauge coupling $\tau^{(0)}_{\text{eff}}$ is given by:

$$\tau^{(0)}_{\text{eff}} \;=\; \tfrac{1}{2}\mathcal{F}^{(0)\prime\prime}(a) \;=\; \tau + \frac{i}{\pi}\sum_{n=0,2,4,\ldots}^{\infty} c_n q^n, \tag{6.97}$$

where q is the exponentiated complexified gauge coupling constant which is the invariant scale of the theory:

$$q \;\equiv\; \exp(i\pi\tau). \tag{6.98}$$

Secondly, when one of the four matter multiplets has a non-zero mass, which we denote m_4, a discrepancy occurs. In the simultaneous scaling limit $m_4 \to 0$ and $g_4 \to 0$, the $N_f = 4$ theory should flow (via renormalization group flow) to the $N_f = 3$ theory. There then exists a relation between the dynamical scale $\Lambda_3^{\rm SW}$ of the $N_f = 3$ theory and the parameters of the $N_f = 4$ theory. Seiberg and Witten give this relation to be:

$$\Lambda_3^{\rm SW} = 64 m_4 \exp\left(-8\pi^2/(g_4^{\rm SW})^2\right), \tag{6.99}$$

where $g_4^{\rm SW}$ is the classical gauge coupling constant in the $N_f = 4$ theory used in Seiberg–Witten theory. In the Pauli–Villars regularization scheme, this formula becomes [219]:

$$\Lambda_3^{\rm SW} = 4\Lambda_3^{\rm PV} = 4 m_4 \exp\left(-8\pi^2/(g_4^{\rm PV})^2\right), \tag{6.100}$$

where $g_4^{\rm PV}$ is the Pauli–Villars gauge coupling of the $N_f = 4$ theory. There is a mismatch between the two formulae for $\Lambda_3^{\rm SW}$ as the numerical constant of proportionality is 64 in Eq. (6.99) and 4 in Eq. (6.100). This discrepancy is resolved by using the non-perturbatively corrected effective complexfied gauge coupling constant $\tau_{\rm eff}^{(0)}$ in Eq. (6.97), for which the appropriate expansion coefficient is $c_0 = 4\log 2$. This then rectifies the discrepancy between the dynamical scale used by Seiberg and Witten and the dynamical scale derived using the Pauli–Villars scale [58] for the $N_f = 3$ theory. If there were no quantum non-perturbative corrections to the effective complexified coupling τ, as given in Eq. (6.97), this discrepancy has no apparent resolution.

The third discrepancy involves the elliptic curve proposed by Seiberg and Witten (Eq. (5.188)) to describe the $N_f = 4$ theory. Due to the discrepancy described above for the $N_f = 3$ theory, the $N_f = 4$ elliptic curve must be corrected if it is to flow (via renormalization group flow) to the corrected $N_f = 3$ theory.

Dorey et. al [219] propose a reparameterized elliptic curve for the $N_f = 4$ theory which agrees with all known perturbative and non-perturbative calculations for this theory. Furthermore, they anticipate the necessity of reparameterizations for the discrepancies between the proposed exact results and instanton calculations in $\mathcal{N} = 2$ $SU(N)$ SQCD with $N_f = 2N$ matter multiplets. Finally, we note that the two-instanton prediction from Seiberg–Witten theory coupled to $N_f = 4$ matter multiplets with arbitrary masses has not yet been determined; the prediction for massless matter multiplets is obtained in [224], which shows that the prepotential in this theory is not equal to its classical value. However, the relation in Eq. (6.97), between the Seiberg–Witten effective coupling $\tau_{\rm eff}^{(0)}$ and the microscopic coupling τ, has been fixed by matching the two-instanton calculation peformed in [217] with the proposed exact results, in [240].

Mass Deformed $\mathcal{N} = 4$ Supersymmetric $SU(2)$ Yang–Mills Gauge Theory

We now briefly describe a discrepancy arising in $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory coupled to one matter multiplet transforming in the adjoint representation of $SU(2)$. If the adjoint matter multiplet is massless in this theory, then the $\mathcal{N} = 2$ supersymmetry is extended to an $\mathcal{N} = 4$ supersymmetry. For this reason, the $\mathcal{N} = 2$ theory coupled to a massless adjoint matter multiplet is referred to as 'mass deformed' $\mathcal{N} = 4$ supersymmetric $SU(2)$ Yang–Mills gauge theory. When there are more than one adjoint matter multiplets coupled to the pure Yang–Mills gauge theory, the theory possesses a positive beta function, and the theory is physically ill defined.

Although this particular model and the associated discrepancy are indirectly related to our description of the matching between proposed exact results and instanton calculations in scale invariant $\mathcal{N} = 2$ $SU(2)$ SQCD, it is instructive to include it since it bears directly on the elliptic curve given in Eq. (5.188) in Section 5.3 of Chapter 5 from which the proposed exact solution for the low energy effective action of this theory derives. In view of our detailed description of Seiberg–Witten theory is Chapter 5, we consider the inclusion of a brief note on this discrepancy relevant.

Dorey et. al [220] report a one-instanton discrepancy in this theory for the instanton contributions to the physical quantum modulus u_2. The proposed resolution for this discrepancy is that the quantum modulus itself receives instanton contributions. However, this proves to be insufficient to rectify the discrepancy. The prediction for the quantum modulus u_2 given by Seiberg and Witten, which we denote by $\tilde{u}_2$, can be extracted from their proposed exact solution of the low energy dynamics of this theory. The general form for the corrected quantum modulus is:

$$\tilde{u}_2 \;=\; u_2 - m^2 \left(\tfrac{1}{12} + \sum_{n=1}^{\infty} \alpha_n q^n \right) , \tag{6.101}$$

where the set $\{\alpha_n\}$ are expansion coefficients, and q is again the exponentiated complexified gauge coupling constant given in Eq. (6.98). The expansion coefficients α_n for the instanton contributions are not predicted by Seiberg–Witten theory, and can only be determined by precise matching with first principles instanton calculations. Note that in the massless limit, $m \to 0$, the $\mathcal{N}=2$ theory becomes the mass deformed $\mathcal{N}=4$ theory, and there are no non-perturbative contributions to $\tilde{u}_2$, as then $\tilde{u}_2 = u_2$.

The one-instanton contribution to u_2 has been calculated in this theory and can be related to the prepotential via the generalized Matone relation [229], given in Eq. (6.81) of Subsection 5.3.2 of Chapter 5, as follows:

$$u_2^{1-\text{inst}}(a) \;=\; 2\pi i \mathcal{F}^{1-\text{inst}}(a) \;=\; \left(\frac{m^4}{a^2} - \frac{m^2}{2}\right) \exp\left(-8\pi^2/g^2\right). \tag{6.102}$$

The prediction of Seiberg and Witten for this quantity in the mass-deformed $\mathcal{N}=4$ theory is given by Eq. (5.189) in Subsection 5.3.2 of Chapter 5, which we reiterate here:

$$\tilde{u}_2 \;=\; u_2 - \tfrac{1}{8} e_1 m^2, \tag{6.103}$$

where e_1 is a root of the elliptic curve for this theory given in Eq. (5.188) of Chapter 5. This prediction corresponds to Eq. (6.101) with exact predictions for all of the expansion coefficients α_n. In particular, one has $\alpha_1 = \alpha_2 = 2$.

However, using the corrected form of the exact quantum modulus $\tilde{u}_2$ derived using Seiberg–Witten theory, in Eq. (6.101), one can derive the one-instanton contribution to u_2 as it should appear in the prediction of Seiberg–Witten theory. This has the form:

$$u_2^{1-\text{inst}} \;=\; \left((2+\alpha_1)m^2 + \frac{m^4}{a^2}\right) \exp\left(-8\pi^2/g^2\right). \tag{6.104}$$

Comparing Eq. (6.102) and Eq. (6.104) gives the result that $\alpha_1 = -5/2$ if the proposed exact result for $u_2^{1-\text{inst}}$ is to match the instanton prediction for the one-instanton contribution. This is in disagreement with the value predicted by Seiberg–Witten theory given in Eq. (6.103), for which one has $\alpha_1 = 2$.

6.5.1 Matching in $\mathcal{N}=2$ Supersymmetric $SU(N)$ Gauge Theory

The forms of the exact solutions proposed for theories with gauge group $SU(N)$ differ from each other, and there exists the serious disagreement between proposed exact results and instanton predictions for $N_f = 2N$ fundamental matter multiplets.

The discrepancies between the instanton predictions and the proposed exact results for $\mathcal{N}=2$ $SU(N)$ SQCD with $N_f = 2N$ matter multiplets have been considered in detail in [219, 223]. In particular, the analysis given in [219] can be extended to the $SU(N)$ theory with $N_f = 2N$. Again, the cause of the mismatch between the results is shown to be a finite perturbative and non-perturbative renormalization of the complexified gauge coupling constants in the low energy effective theory. In order to match the two sets of results exactly, this renormalization of the coupling constants must be taken into account. This shows that explicit instanton corrections are required to fix the proposed exact results in this case. The parameterization of the quantum moduli space used in the generalized Seiberg–Witten theory is not equivalent to the parameterization given by the gauge invariant quantum moduli u_n. Fixing the two sets of results has been achieved in special cases at the one-instanton level in [188, 215, 218, 219, 223, 221]. A general matching prescription to resolve these discrepancies was proposed in [188]. This over-arching scheme, referred to as the Argyres–Pelland matching scheme, claims to generalize the proposals for matching the two sets of results above following from one-instanton and two-instanton checks. It is designed for the matching of proposed exact results and instanton calculations in $\mathcal{N} = 2$ $SU(N)$ SQCD with N_f massive or massless fundamental matter multiplets. The Argyres–Pelland scheme also purports to resolve the differences between the hyperelliptic curves proposed to describe the quantum moduli space in this theory. This is done by introducing a general form of hyperelliptic curve which is claimed to correctly describe the quantum $SU(N)$ SQCD moduli space, and from which all of the previously proposed hyperelliptic curves for this theory can be recovered.

By considering the permissible non-perturbative redefinitions of the physical quantities involved, Argyres and Pelland are able to eliminate the reported ambiguities between the proposed exact solutions themselves and between the results from these and their instanton counterparts. Hence the exact results and instanton predictions are in agreement modulo the permitted reparameterizations. The constants in the matching scheme can only be fixed by comparison with instanton calculations, however, and cannot be derived from the exact results themselves. This necessitates use of the predictions for the instanton contributions to the prepotential stated in Section 6.3.

In this subsection we perform a non-trivial test of the matching scheme in [188] for the entire quantum moduli space in $SU(N)$ theories with $N_f = 2N$ massless fundamental hypermultiplets. This is done by determining the constants which give agreement between the proposed exact results and the one-instanton predictions for general $N > 1$. Previously, this one-instanton check had been performed for a single special point of the moduli space in these theories [188], where agreement was found within the scheme.

The Argryes–Pelland Matching Scheme

We begin by briefly reviewing the Argyres–Pelland matching scheme [188] and also the conjectured method of exact solution for $\mathcal{N} = 2$ supersymmetric $SU(N)$ gauge theories. The gauge coupling parameter used, valid near weak coupling, is the exponentiated complexified gauge coupling constant $q \equiv \exp(2\pi i\tau) \in \mathbb{C}$, which characterizes the scale invariant theory, where:

$$\tau = \frac{\vartheta}{2\pi} + \frac{4\pi i}{g^2} \tag{6.105}$$

is the complexified gauge coupling constant, for gauge coupling g. In the weak coupling regime, $q \approx 0$.

The matching scheme is derived by considering the most general mapping between the parameters and scalar vacuum expectation values of each set of results. These must be consistent at weak coupling and obey the constraints imposed by supersymmetry [188]. Furthermore, the matching scheme agrees with dimensional analysis and also ensures that the notion of the moduli space is preserved.

For $\mathcal{N} = 2$ supersymmetric $SU(N)$ theories, the low energy effective action is a function of the bare masses $\{m_n\}$, $n = 1 \ldots 2N$, the coupling parameter q, and the set of scalar vacuum expectation values, or moduli, is $\{u_i\}$, $i = 1 \ldots N$, all of which assume complex values. We denote the parameters and scalar vacuum expectation values appearing in the proposed exact solutions of these theories as the set $\{\tilde{q}, \widetilde{m}_n, \tilde{u}_i\}$. The counterparts of these parameters appearing in the instanton predictions are denoted by $\{q, m_n, u_i\}$. The constants in the matching scheme are denoted by $\{C_s, B_s, A_s^{(i;\{i_m\})}\}$. Holomorphy and the asymptotic behaviour at weak coupling imply that the general map between q and $\tilde{q}$ is given by:

$$\tilde{q} = \sum_{s=0}^{\infty} C_s q^{s+1}. \tag{6.106}$$

Dimensionless ratios of the masses cannot enter into Eq. (6.106) due to the matching that must hold at very weak coupling. The masses $\{m_n\}$ and $\{\widetilde{m}_n\}$ are related via:

$$\widetilde{m}_n = \left(1 + \sum_{s=1}^{\infty} B_s q^s\right) m_n. \tag{6.107}$$

Finally, the matching relation between the scalar vacuum expectation values $\{u_i\}$ and $\{\tilde{u}_i\}$ is given by:

$$\tilde{u}_i = \sum_{s=0}^{\infty} A_s^{(i;\{i_m\})} q^s \prod_{m=0}^{N} u_{i_m}, \tag{6.108}$$

in which we define $u_0 \equiv m$ following Argyres and Pelland [188].

One-Instanton Level Matching of Instanton Predictions and Exact Results

We shall employ the relations in Eqs. (6.106, 6.108) to one-instanton level, or, equivalently, to $\mathcal{O}(q)$, in a test of the Argyres–Pelland matching scheme. The matching in Eq. (6.107) between the masses is not required here since we consider all fundamental matter multiplets in the theory to possess zero mass.

The defining quantity in $\mathcal{N} = 2$ supersymmetric $SU(N)$ theories is the prepotential, $\mathcal{F}$, a function of the superfields, which determines the low-energy effective (Wilsonian) action of the theory.

The prepotential can be decomposed into a classical part ($\mathcal{F}_{\text{cl}}$) and its perturbative corrections ($\mathcal{F}_{1-\text{loop}}$), which are one-loop exact in this case (due to nonrenormalization theorems), and non-perturbative corrections, ($\mathcal{F}_{\text{inst}}$), containing instanton effects of all orders, as given in Eq. (5.42) in Section 5.3 of Chapter 5.

The instanton contributions in this particular theory have the form of an infinite sum involving powers of the gauge coupling parameter q:

$$\mathcal{F}_{\text{inst}} = \sum_{k=0}^{\infty} q^k \mathcal{F}_k, \tag{6.109}$$

where $\mathcal{F}_k = \mathcal{F}_k(a_i)$ are functions of the classical vacuum expectation values, $\{a_i\} \in \mathbb{C}$, of the scalar superfield ϕ in the adjoint representation (i.e., the Higgs field), and k is the instanton number or charge. All charges of instantons contribute in this case as the $\mathbb{Z}_2$ parity symmetry of the case when only one multiplet is massless is absent.

The superfield ϕ is a member of the vector multiplet of the theory. In general, the prepotential $\mathcal{F}$ is a holomorphic function of the scalar vacuum expectation values $\{a_i\}$. In the scalar invariant theories considered here, the perturbative beta function vanishes, and so the perturbative corrections vanish: $\mathcal{F}_{1-\text{loop}} = 0$. The scalar vacuum expectation values of the electric-magnetic dual of ϕ are $\{a_{Di}\} \in \mathbb{C}$, and are related to the prepotential $\mathcal{F}$ via:

$$a_{Di} = \frac{\partial \mathcal{F}}{\partial a_i}. \tag{6.110}$$

The gauge coupling matrix of the theory is given by:

$$\tau_{ij} = \frac{\partial^2 \mathcal{F}}{\partial a_i \partial a_j}. \tag{6.111}$$

The scalar potential of the $\mathcal{N} = 2$ supersymmetric Lagrangian describing the vector multiplet is a function of ϕ. At weak coupling, the vacuum expectation value of the scalar field ϕ is given by the matrix:

$$\langle \phi \rangle = \text{diag}[a_1, \ldots, a_N]. \tag{6.112}$$

The region of the quantum moduli space corresponding to non-zero vacuum expectation values of the scalar fields in the vector multiplet of the theory is referred to as the Coulomb branch. In this phase of the theory only the scalar field ϕ acquires a vacuum expectation value. The holomorphic co-ordinate on the classical moduli space is a function of the classical vacuum expectation values and the field ϕ, which we denote by $\{u_i\} \in \mathbb{C}$, $i = 1, \ldots, N$, and refer to as the classical moduli. These moduli can be written in terms of the classical scalar vacuum expectation values $\{a_i\}$ of ϕ as follows:

$$u_n^{\text{cl}} = \langle \text{tr}(\phi^n) \rangle = (-1)^n \sum_{i_1 < \ldots < i_N}^{N} a_{i_1} \cdots a_{i_N}. \tag{6.113}$$

For $SU(N)$ theories, classically one has $\text{tr}(\phi) = \sum_{i=1}^{N} a_i = 0$, and $u_1 = 0$, by definition. In particular, the first non-zero classical moduli is given by:

$$u_2^{\text{cl}} = \langle \text{tr}(\phi^2) \rangle = \frac{1}{2} \sum_{i=1}^{N} a_i^2. \tag{6.114}$$

The quantum corrections to the classical modulus, or condensate, u_2^{cl}, will be the focus of the one-instanton matching we determine in this theory.

The classical moduli space of these theories does not receive quantum corrections, as it is protected by $\mathcal{N}=2$ supersymmetry, but the metric ($ds^2 = \mathrm{Im}\,\tau_{ij}\,\mathrm{d}a_i \mathrm{d}\bar{a}_j$) on it does. Thus the classical scalar vacuum expectation values $\{a_i\}$ receive quantum corrections, which then in turn determine the quantum corrections to the classical moduli $\{u_i\}$. Following the methods introduced by Seiberg and Witten in [170, 171], which were described in detail in Section 5.3 of Chapter 5, one identifies the $(N-1)$-dimensional quantum moduli space with the moduli space of a genus $(N-1)$ compact Riemann surface. Then the functions $\{a_i, a_{Di}\}$ can be calculated as the periods, about certain cycles, of the Riemann surface, and the gauge coupling matrix τ_{ij} (Eq. (6.111)) is the period matrix of such a surface.

A standard result of the theory of algebraic curves [249, 250] is that any compact Riemann surface can be specified completely by a class of elliptic $(N-1 \leq 2)$ or hyperelliptic $(N-1 > 2)$ curves. For the moduli spaces considered here, these curves have the form:

$$y^2 \; = \; F^2(x) - G(x), \tag{6.115}$$

where $F(x)$ is a polynomial of degree $(N-1)$ in the dummy variable $x \in \mathbb{C}$, whose coefficients are functions of the set of moduli $\{u_i\}$. The roots of $F(x)=0$ are the exact vacuum expectation values $\{e_i\}$, which parameterize the quantum moduli space, and which are related to the classical scalar vacuum expectation values $\{a_i\}$ by:

$$a_n \; = \; (-1)^n \sum_{i_1<\ldots<i_N} e_{i_1}\cdots e_{i_N}. \tag{6.116}$$

These parameters obey $\sum_{i=1}^{N} e_i = 0$.

The first non-zero quantum moduli (i.e. moduli of the quantum moduli space) in the exact results proposed for this theory is then:

$$\tilde{u}_2^{\mathrm{qu}} \; = \; \frac{1}{2}\sum_{i=1}^{N} e_i^2. \tag{6.117}$$

In our conventions the hyperelliptic curve associated with the matching prescription for $SU(N)$ theories with $N_f = 2N$ massless fundamental matter hypermultiplets is given by Argyres and Pelland [188] to be:

$$y^2 \; = \; \left(x^N - \sum_{k=1}^{N-1} \tilde{u}_{k+1} x^{N-k-1}\right)^2 - \tilde{q}x^{2N}. \tag{6.118}$$

The functions $\{a_i\}$ and $\{a_{Di}\}$ can be determined by evaluating the meromorphic one-form λ, where:

$$\lambda = \frac{xdx}{2\pi i y}\left[\frac{F(x)G'(x)}{2G(x)} - F'(x)\right], \tag{6.119}$$

in which the prime denotes differentiation with respect to x, over the canonical basis of homology one-cycles $\{\alpha_i, \beta_i\}$ of the Riemann surface:

$$a_i = \oint_{\alpha_i} \lambda, \tag{6.120}$$

$$a_{Di} = \oint_{\beta_i} \lambda, \tag{6.121}$$

which is a special case of Eq. (5.212) of Subsection 5.4.1 of Chapter 5.

Given the curve defining the moduli space of the theory, one can then exactly determine the quantum moduli $\{u_i^{\rm qu}\}$, and hence the prepotential $\mathcal{F}$, via Eqs. (6.110, 6.119, 6.120, 6.121), following the method of Seiberg and Witten described in Section 5.3 of Chapter 5.

One can perform the integration of the meromorphic one-form in Eq. (6.119) exactly for $SU(N)$ theories. This has been done previously in Ref. [218]. Using the curve in Eq. (6.118), to order $\mathcal{O}(\tilde{q})$ (where we use $\mathcal{O}(x)$ to denote the order of the variable x) one has:

$$a_i = \oint_{\alpha_i} \frac{dx}{2\pi i}\left(N - \frac{xF'(x)}{F(x)} + \tilde{q}\frac{x^{2N}(NF - xF')}{2F(x)^3} + \mathcal{O}(\tilde{q}^2)\right). \tag{6.122}$$

At weak coupling, the homology one-cycles $\{\alpha_i\}$ coincide with the exact quantum vacuum expectation values $\{e_i\}$. We note that the electric-magnetic duality ambiguity [188] in the case considered here is trivial. Discarding a total derivative and performing the integration in Eq. (6.122) yields the following $(N-1)$ equations:

$$a_i = e_i + \tilde{q}\frac{e_i^{2N-1}}{2\Delta_i(e_i)}\left(N - \sum_{i\neq j}\frac{e_i}{(e_i - e_j)}\right), \tag{6.123}$$

where $\Delta(e_i) = \prod_{i\neq l}(e_i - e_l)$, $l = 1 \ldots N-1$. In obtaining this result, the reverse direction was taken in performing the period integral Eq. (6.120), so that the classical scalar vacuum expectation values $\{a_i\}$ remain positive. Solving Eq. (6.123) at leading order in $\tilde{q} \ll 1$ yields the exact vacuum expectation values $\{e_i\}$:

$$e_i = a_i - \tilde{q}\frac{a_i^{2N-1}}{2\Delta_i(a_i)}\left(N - \sum_{i\neq j}\frac{a_i}{(a_i - a_j)}\right). \tag{6.124}$$

Equations (6.113, 6.116, 6.124) show that at the classical level the expected results are reproduced.

We now write $\Delta_i \equiv \Delta(a_i)$ for simplicity. Using the definition in Eq. (6.117), the exact result $\tilde{u}_2^{\mathrm{qu}}$ is given by:

$$\tilde{u}_2^{\mathrm{qu}} = u_2^{\mathrm{cl}} + \tilde{q}\tilde{u}_2^{\mathrm{inst}} = \frac{1}{2}\sum_{i=1}^{N} a_i^2 - \tilde{q}\sum_{i=1}^{N}\frac{a_i^{2N}}{2\Delta_i^2}\left(1 - \sum_{i\neq j}\frac{a_j}{(a_i - a_j)}\right). \tag{6.125}$$

Equation (6.125) gives the explicit form of $\tilde{u}_2$, which is the value of the quantum modulus u_2 predicted by the exact results based on the methods of Seiberg–Witten theory. This expression for the exact quantum modulus $\tilde{u}_2$ holds for general values of $N > 1$ and agrees up to regular terms with the previous results given in [218]. These regular terms are non-singular terms dependent upon the exact quantum vacuum expectation values $\{e_i\}$. The generalized Matone relation [229], given by Eq. (6.82) of Section 6.3, enables one to relate the exact quantum modulus $\tilde{u}_2^{\mathrm{qu}}$ to the one-instanton prepotential $\mathcal{F}_1|_{N,N_f}$ derived from the exact results, and we shall employ this in a test of the Argyres–Pelland matching scheme.

The Argyres–Pelland matching scheme [188] purports to account for the most general mapping which can connect the parameters and the moduli for both sets of results. The generic parameters and moduli are $\{\tilde{q}, \tilde{u_i^{\mathrm{qu}}}\}$ for the proposed exact results, and $\{q, u_i^{\mathrm{qu}}\}$ for the instanton results, since the matter multiplet masses are set to zero here and do not enter into the calculation here. The Argyres–Pelland matching prescription to $\mathcal{O}(q)$ for the coupling parameter q and the quantum moduli u_2^{qu} is given by the following set of relations:

$$\tilde{q} = C_0 q, \tag{6.126}$$

$$\tilde{u}_2^{\mathrm{qu}} = (1 + A_1^{(1;1)} q) u_2^{\mathrm{qu}}. \tag{6.127}$$

In the Argyres–Pelland matching scheme, no modular invariance (S-duality) is assumed for the space of couplings in the reparameterization in Eqs. (6.106, 6.126), as it is not necessarily a physical attribute of the $N_f = 2N$ theory. We now need to isolate the non-perturbative contribution to the classical vacuum expectation value u_2^{cl}, as it is this component which contains the instanton corrections. Writing Eqs. (6.126, 6.127) in terms of perturbative and non-perturbative parts gives:

$$C_0 \tilde{u}_2^{\text{inst}} \;=\; u_2^{\text{inst}} + A_1^{(1;1)} u_2^{\text{cl}}. \tag{6.128}$$

The classical modulus u_2^{cl} in the matching relation above (Eq. (6.128)) constitutes a regular term; the other terms will then have the same singularity structure, according to [218].

The generalized Matone relation [229] for the quantities found using instanton calculus has the form given in Eq. (6.82) of Section 6.3 in Chapter 5. The result for the one-instanton prepotential $\mathcal{F}_1|_{N,N_f}$ for $\mathcal{N}=2$ supersymmetric $SU(N)$ Yang–Mills gauge field theories with $N_f = 2N$ massless fundamental matter multiplets [223] is given by Eq. (6.94) of Section 6.3. For the case of $N_f = 2N$ massless fundamental matter multiplets, one has $\{m_i\} = 0$, and Eq. (6.94) gives [188]:

$$\frac{1}{2\pi i} u_2^{\text{inst}} \;=\; \mathcal{F}_1|_{N,N_f} \;=\; -\frac{i C_1' \pi^{2N-1}}{2^{2N+2}} \sum_{i \neq j} \frac{(a_i + a_j)^{2N}}{(a_i - a_j)^2 \gamma_i \gamma_j}, \tag{6.129}$$

where $\Delta_i = (a_i - a_j)\gamma_i$ and $\gamma_i = \prod_{i\neq k, k\neq j}(a_i - a_k)$, $k = 1 \ldots N-2$. The constant C_1' is the renormalization scheme-dependent one-instanton factor, also known as the 't Hooft one-instanton factor [18]. Following Argyres and Pelland [188] we use the standard value of C_1' given by:

$$C_1' \;=\; 2^{N+2} \pi^{-2N}. \tag{6.130}$$

Inserting Eqs. (6.125, 6.129) and the exact result for $\tilde{u}_2^{\text{qu}}$, from Eq. (6.125), into Eq. (6.128), we derive the following relation which matches the instanton contributions to the prepotential and the proposed exact results within the Argyres–Pelland scheme:

$$-C_0 \sum_{i=1}^{N} \frac{a_i^{2N}}{2\Delta_i^2} \left(1 - \sum_{i \neq j} \frac{a_j}{a_i - a_j} \right) \;=\; \frac{C_1' \pi^{2N}}{2^{2N+1}} \sum_{i=1}^{N} \sum_{i \neq j} \frac{(a_i + a_j)^{2N}}{(a_i - a_j)^2 \gamma_i \gamma_j} + \frac{1}{2} A_1^{(1;1)} \sum_{i=1}^{N} a_i^2. \tag{6.131}$$

To extract the one-instanton matching coefficient C_0, one observes that manipulating Eq. (6.131) so that both sides of the equality have the same denominator enables one to take the previously singular limit $a_i \to a_j$. This manipulation involves the following equality [188]:

$$\sum_{i}^{N} \sum_{i \neq j} \frac{(a_i + a_j)^{2N}}{(a_i - a_j)^2 \gamma_i \gamma_j} \;=\; \sum_{i} 2^{2N} \frac{a_i^{2N}}{\Delta_i^2} - \sum_{i}^{N} \sum_{j}^{N} \frac{(a_i + a_j)^{2N}}{\Delta_i \Delta_j}. \tag{6.132}$$

which can be used to write both sides of Eq. (6.131) in terms of fractions with denominator Δ_i^3. After some cancellations, the singular limit $a_i \to a_j$ can then be taken, and the coefficients of the non-vanishing terms in the expression can be compared. Via this, one then obtains C_0 as:

$$C_0 = 2^{N+2}. \tag{6.133}$$

The form of C_0 found in Eq. (6.133) is in exact agreement with the form of C_0 determined in [188] for a single point of the moduli space in $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge field theories.

Using the expression for C_0 given above, one can explicitly determine the constant $A^{(1;1)}$ by expanding Eq. (6.131) and comparing the coefficients of the leading order terms. To do this, we expand the principal objects in Eq. (6.129) as:

$$\Delta_i \approx a_i^{N-1} + \dots, \tag{6.134}$$

$$\gamma_i \approx a_i^{N-2} + \dots, \tag{6.135}$$

$$(a_i + a_j)^{2N} = \sum_{r=0}^{2N} \begin{pmatrix} 2N \\ r \end{pmatrix} a_i^{2N-r} a_j^r, \tag{6.136}$$

and compare the coefficients of the terms of highest order in a_i in Eq. (6.131). This then implies the following value for the remaining one-instanton matching coefficient, $A_1^{(1;1)}$:

$$A_1^{(1;1)} = -2^{N+2} + 2^{2-N} \begin{pmatrix} 2N \\ N-1 \end{pmatrix}. \tag{6.137}$$

The formulae Eqs. (6.133, 6.137) are valid for general values of $N > 1$. This can be checked by an inductive argument using Eq. (6.131) and Eq. (6.132).

It has been shown that the form of the Argyres–Pelland matching relations between the exact results and the instanton predictions is correct at the one-instanton level, since we have obtained agreement between the instanton predictions and the proposed exact results, for the quantum modulus u_2^{qu}, for all $N > 1$.

We now comment on its relevance to the string theoretic derivation of the class of hyperelliptic curves corresponding to those found in the exact solutions [274]. The classical brane configuration of N D4-branes suspended between two parallel NS5-branes in Type-IIA string theory corresponds to the vacua of classical $\mathcal{N} = 2$ supersymmetric Yang-Mills theory, and the vacua of the quantum theory corresponds to the supersymmetric configurations of an M-theory $M5$-brane with a particular world volume [274]. To incorporate N_f matter hypermultiplets into the system, one attaches N_f semi-infinite 4-branes to the NS5-branes. The class of curves corresponding to those appearing in the proposed exact results follows from this brane configuration. For detailed reviews of this construction, see, for example, References [275, 276].

The dictionary [274] set up between the parameters of the brane configuration and the parameters of the field theory is only valid at extremely weak coupling. Beyond extremely weak coupling, quantum corrections will in general modify this dictionary, and it will contain ambiguities manifest as the freedom to make non-perturbative redefinitions of the parameters. Hence, the brane–field theory correspondence is valid, but the quantitative dictionary connecting them is ambiguous.

The Argyres–Pelland matching scheme [188] uses the most general permissible redefinitions of the parameters and the vacuum expectation values of these field theories. It is natural to propose that the ambiguities in the M-theoretic derivation of the curves which exactly solve the low-energy effective actions of the same field theories are resolved by the same matching scheme. That is, the equivalence class of curves derived from M-theory should coincide precisely with the equivalence class of curves derived from the exact solutions. Then the mappings between the elements of the equivalence class of M-theoretic curves will be the same mappings between elements of the equivalence class of exact solution curves. Hence the same matching scheme, namely the Argyres–Pelland matching scheme, for comparing the exact results to instanton results should also hold for comparing the M-theoretic results to instanton results. A precise test of this conjecture in the context of string theory would validate the Argyres–Pelland matching scheme for future M-theory predictions.

To conclude this subsection, we have found that the proposed exact results and the instanton predictions can be matched to one-instanton level for $\mathcal{N} = 2$ supersymmetric gauge theories with gauge group $SU(N)$ and $N_f = 2N$ massless fundamental matter multiplets, for general $N > 1$ within the Argyres–Pelland matching scheme. In particular, this matching was explicitly performed for the proposed exact results pertaining to the first non-zero quantum modulus u_2^{qu}. The coefficients which implement this matching are the constants C_0 and $A_1^{(1;1)}$, given by Eqs. (6.133, 6.137). It has been shown that this matching can be achieved for the complete quantum moduli space of these theories at the one-instanton level, extending the previous result of Argyres and Pelland [188], whose results held for a moduli subspace.

The case where the $N_f = 2N$ hypermultiplets have non-zero arbitrary masses could also be investigated; one expects that the constants C_0 and $A_1^{(1;1)}$ in the massive case should agree with the those above when all N_f multiplet masses are set to zero. This is in accord with renormalization group flow arguments. This task is more complicated than the above matching in the theory with massless multiplets as the matching relation for the masses Eq. (6.107) is then required. It was found in a preliminary investigation that the introduction of non-zero multiplet masses poses hitherto unexpected difficulties in attempting to use the Argyres–Pelland matching relations. However, we believe that these are technical problems which can be overcome and that the general principles on which the Argyres–Pelland matching scheme are based are valid and correct.

Tests of the Argyres–Pelland scheme at the two-instanton level in $\mathcal{N} = 2$ $SU(N)$ SQCD with $N_f \leq 2N$ fundamental matter multiplets would be desirable since these would provide a physical check of the Argyres–Pelland matching beyond the above constraints on the values of the constants involved.

Chapter 7

Conclusion

In this thesis we have we have presented results concerning exact non-perturbative physics in globally supersymmetric quantum field theory and exact classical solutions of self-dual Yang–Mills equations. The former concerns the matching of instanton predictions and proposed exact results in scale invariant $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD. The latter concerns the general $U(N)$ two-instanton solution determined using the ADHM construction of instantons. Our results were detailed in Chapter 2 and Chapter 6.

In Chapter 2 we described the phenomenon of instantons in Yang–Mills gauge field theories. The first instanton configuration found, the BPST instanton, was described. Generic properties of instantons were then described, and the important concept of the instanton moduli space was outlined. The ADHM construction of instanton solutions, which implicitly defines all exact instanton gauge field configurations with classical gauge group and arbitrary topological charge k, was described in detail for the gauge group $U(N)$. Using the ADHM construction, the exact general $U(N)$ ADHM two-instanton solution was explicitly given. This was the result of studying the ADHM instanton constraints for gauge group $U(N)$ and topological charge $k = 2$. An explicit and general solution, valid for all $N > 1$, was determined using linear algebra. This was done by exploiting the linearity present within the constraints. The construction of the $U(N)$ two-instanton gauge field following from this explicit configuration was also outlined. The instanton gauge field configuration found constitutes an exact general solution of the self-dual Yang–Mills field equations with gauge group $U(N)$ or $SU(N)$ and topological charge (or instanton number) $k = 2$. Hence it is an exact solution of the classical equations of motion of the $U(N)$ Yang–Mills action in four dimensional Euclidean spacetime which gives the minimum finite value of the Yang–Mills action (in the $k = 2$ sector of topological charge). This result gives the first exact general $U(N)$ ADHM multi-instanton configuration. The solution may be used in semi-classical calculations of non-perturbative effects in $U(N)$ and $SU(N)$ Yang–Mills gauge field theories, via the collective co-ordinate method.

The ADHM constraints for gauge group $U(N)$ and topological charge $k = 3$ were described in detail. The method of solution for the $U(N)$ $k = 2$ constraints did not apply in this case as the $k = 3$ constraints possess a greater degree of non-linearity, are more highly coupled (with bilinear terms), and there are a greater number of constraints. There appears to be no underlying principles which could assist in the solution of these constraints for $k = 3$ and $k \geq 4$. We also note that the $U(N)$ ADHM constraints for topological charge $k \geq 4$ are ambigious in the allocation of physically identifiable parameters.

The ADHM construction for Yang–Mills instanton gauge fields with gauge group $Sp(N)$ was also described in Chapter 2. The ADHM construction for the symplectic groups requires fewer variables and generates a smaller number of constraints than the $U(N)$ ADHM formalism. The $Sp(N)$ ADHM construction is useful for constructing instantons with the simplest non-trivial non-Abelian gauge group, $SU(2)$, through the isomorphism $Sp(1) \simeq SU(2)$. The exact general $Sp(N)$ two-instanton is the only exact and general $Sp(N)$ multi-instanton configuration known. The $Sp(N)$ ADHM three-instanton constraints were then described. These constraints form a set of simultaneous non-linear quaternionic equations. We were not able to determine the general solution of these constraints, but described the two existing special exact solutions and also made conjectures regarding the properties of the exact general $Sp(N)$ three-instanton configuration.

The motivation for the determination of the instanton result in Chapter 2 was testing the proposed exact results in four dimensional $\mathcal{N} = 2$ supersymmetric quantum Yang–Mills gauge field theories. Supersymmetric gauge theories are field theories which possess a special symmetry which rotates fermionic and bosonic states into one another, known as global supersymmetry, under which the theory is invariant. In Chapter 3 we described supersymmetric gauge theories, beginning with the supersymmetry algebra and fundamental aspects of global supersymmetry. We followed this with a brief description of supersymmetry constraints. The superfield formalism for supersymmetric field theories was then described. Through this formalism, one can construct $\mathcal{N}$-extended supersymmetric field theories from $\mathcal{N} = 1$ supersymmetric fields. We then briefly described each of the $\mathcal{N}$-extended supersymmetric gauge theories in turn, up to the case of maximally extended supersymmetric gauge theories, for which $\mathcal{N} = 4$. The most interesting supersymmetric gauge theories from a physical and phenomenological perspective are the $\mathcal{N} = 1$, $\mathcal{N} = 2$ and $\mathcal{N} = 4$ supersymmetric gauge theories. In this review, preliminaries for the next two chapters were given. For completeness, we noted that field theories with $\mathcal{N} = 3$ supersymmetry are intrinsically unphysical and are difficult to describe.

In Chapter 4 we described some of the exact results which have been obtained for field theories with $\mathcal{N} = 1$ and $\mathcal{N} = 4$ supersymmetry. These include results which are generic properties of supersymmetric gauge theories, and other results which have only been found for these particular theories. Instanton calculations have proven useful in extracting the exact form of the beta function for these theories. Later in Chapter 4 we outlined the Montonen–Olive conjecture, and the more general form of electric-magnetic duality, known as S-duality, conjectured to exist in $\mathcal{N} = 4$ supersymmetric gauge theories. This required a general but brief description of magnetic monopoles in gauge theories. We then described the modular invariance, or S-duality, which $\mathcal{N} = 4$ supersymmetric Yang–Mills gauge theory is conjectured to possess, and some of the evidence supporting this conjecture. The concept of electric-magnetic duality in gauge theories is central to the proposed exact results in $\mathcal{N} = 2$ supersymmetric gauge theories later described in Chapter 5. Further to this review, we also described Seiberg duality and other forms of duality in $\mathcal{N} = 1$ supersymmetric gauge theories. The notion of a moduli space of vacua and phases of field theories were also described in this chapter, which are important features of $\mathcal{N} = 1$ and $\mathcal{N} = 2$ supersymmetric gauge theories.

The exact results described in Chapter 4 are both complementary and in some cases analogous to the description of the exact results proposed for $\mathcal{N} = 2$ supersymmetric gauge theories in Chapter 5. There exist exact results in $\mathcal{N} = 2$ supersymmetric gauge theories which also use instanton calculations and other methods; some of these are briefly described in Chapter 4. The exact results in $\mathcal{N} = 2$ supersymmetric gauge theories which are the focus of this thesis are those proposed by Seiberg and Witten. In Chapter 5 we described the proposed determination of the low energy Wilsonian effective action of four dimensional $\mathcal{N} = 2$ supersymmetric $SU(2)$ quantum Yang–Mills gauge field theory, known as Seiberg–Witten theory. This is the first known exact solution for the strongly coupled dynamics of a four dimensional quantum field theory. The quantum moduli space of vacua of this theory is claimed to be exactly described by a particular Riemann surface, which is parameterized by a family of elliptic curves. We described the pioneering techniques used to determine these auxiliary elliptic curves, some of which are similar to the results for $\mathcal{N} = 1$ supersymmetric gauge theories described in Chapter 4. The exact low energy effective Wilsonian action for $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory coupled to N_f fundamental matter multiplets, also known as $\mathcal{N} = 2$ supersymmetric $SU(2)$ QCD, which has an exact form also proposed by Seiberg and Witten, was also described. We then described the generalizations of Seiberg–Witten theory to other gauge groups and matter multiplets, with special emphasis upon $\mathcal{N} = 2$ supersymmetric $SU(N)$ QCD, which is the field theory of primary interest in this thesis. Using Seiberg–Witten methods, the moduli spaces of these theories are generically described by families of auxiliary hyperelliptic curves.

In Chapter 6 we described the tests of the proposed exact results in $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge field theories, and primarily those which use instanton calculations. The non-perturbative predictions of the Seiberg–Witten solution can be compared with appropriate instanton calculations, enabling one to test, check and match the exact results proposed by Seiberg and Witten and generalized by others. To perform such instanton calculations requires an instanton calculus, which is a body of methods and results with which quantitative predictions for the non-perturbative quantum corrections arising from instantons can be determined. The semi-classical approximation is made in order to extract such results, and the collective co-ordinate method is employed to permit calculations to be performed. A comprehensive instanton calculus for supersymmetric gauge theories has been developed, and with it sophisticated calculations can be performed which yield exact field theoretic results for the instanton contributions to potentials and Green's functions within $\mathcal{N} = 2$ supersymmetric Yang–Mills gauge field theories. The supersymmetric instanton calculus can be used in conjunction with the ADHM construction of instantons to obtain these results. In Chapter 6 we described the instanton calculations which have been used to test Seiberg–Witten theory and the proposed exact results in $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge field theories. For $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with $N_f \leq 2N - 1$ fundamental matter multiplets, which includes the arena of Seiberg–Witten theory as the case $N = 2$, precise quantitative agreement between the proposed exact results and instanton calculations exists.

For the finite scale invariant theories $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory with $N_f = 4$ fundamental matter multiplets, and $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theories with $N_f = 2N$ fundamental matter multiplets in general, there exist discrepancies between the non-perturbative predictions made using Seiberg–Witten methods and the results of instanton calculations. We reported on these discrepancies and described the first suggested resolution of these discrepancies, made for the case of $\mathcal{N} = 2$ supersymmetric $SU(2)$ Yang–Mills gauge theory with $N_f = 4$ fundamental matter multiplets. A non-perturbative renormalization of the quantum modulus u_2^{qu}, a holomorphic co-ordinate used to parameterize the quantum moduli space of vacua, and the complexified gauge coupling constant τ has been proposed to eliminate the discrepancies. Using a generalization of this work, a scheme known as the Argyres–Pelland reparameterization scheme, which purports to match the two sets of independently derived exact results for $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with $N_f = 2N$ fundamental matter multiplets is then examined. Working within the Argyres–Pelland scheme, we were able to precisely match the one-instanton contribution to the prepotential $\mathcal{F}$ of low energy effective $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with $N_f = 2N$ massless fundamental matter multiplets with the prediction derived from the proposed exact low energy solution of this theory, which is given by a particular hyperelliptic curve. This matching is valid for all values of $N > 1$. In the case of $N_f = 2N$ massive fundamental matter multiplets, unanticipated difficulties were encountered for this matching, and we were not able to generalize our result to the massive case. In Chapter 2 we also commented on the feasibility of using the exact general $U(N)$ ADHM two-instanton configuration, presented in Chapter 2, in extending this matching to the two-instanton level. This would require using the exact general $U(N)$ ADHM two-instanton configuration within the $\mathcal{N} = 2$ supersymmetric instanton calculus to extract the two-instanton contribution to the $SU(N)$ prepotential $\mathcal{F}$. In commutative spacetime, this calculation is expected to be a technically challenging one which would involve further development of analytical methods for semi-classical instanton calculations.

In conclusion, we have undertaken work which aims to quantitatively resolve the discrepancies between the exact non-perturbative results proposed for $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with $N_f = 2N$ fundamental matter multiplets and predictions for the same theory made using instanton calculus. As part of this programme of testing and matching the results of Seiberg–Witten theory and its generalization through instanton calculations, we have explicitly obtained the first known exact general multi-instanton gauge field configuration for the gauge group $U(N)$. This is the exact general $U(N)$ two-instanton field configuration, which was determined using the ADHM construction of instantons. This instanton configuration can potentially be used to calculate the two-instanton contribution to the prepotential of $\mathcal{N} = 2$ supersymmetric $SU(N)$ Yang–Mills gauge theory with N_f fundamental matter multiplets. However, this calculation will involve overcoming the technical difficulties associated with the supersymmetric instanton calculus for this case. The exact general $U(N)$ ADHM two-instanton presented here can also be used in other field theoretic applications. We also considered the $U(N)$ ADHM three-instanton constraints, but were not able to extend the method used to solve the $U(N)$ ADHM two-instanton constraints to this case. The method for solving the $U(N)$ ADHM two-instanton constraints may assist in uncovering other multi-instanton configurations in future work.

The exact general $SU(2)$ three-instanton solution was also implicitly investigated using the ADHM construction. If the exact general $SU(2)$ ADHM three-instanton can be determined, it also can potentially be used in a further multi-instanton test of Seiberg–Witten theory. The observations made here towards a determination of the exact general $SU(2)$

three-instanton may also assist in future work in this direction.

We remain optimistic about future investigations involving instantons and their physical effects in quantum field theory. Supersymmetric field theory has provided a wealth of results which not only may illuminate methods and approaches for exact results in phenomenological quantum field theories, but which also may prove important in the mathematical description of Nature itself. Instantons are physical phenomenon which are likely to play an important rôle in future non-perturbative quantum physics, and whose complete description requires further elucidation.

Appendix A

Conventions

Where stated, and for most of this thesis, we work in four dimensional Minkowski spacetime, using the conventional Minkowski spacetime metric, as in [122], given by:

$$\eta_{mn} = \mathrm{diag}(-1,1,1,1). \tag{A.1}$$

Euclidean Spacetime

When not using Minkowski spacetime, we work in Euclidean spacetime. The metric of four dimensional Euclidean spacetime is taken to be:

$$\eta_{mn} = \mathrm{diag}(1,1,1,1), \tag{A.2}$$

which results from the continuation of four dimenstional Minkowski spacetime to four dimensional Euclidean spacetime. This essentially replaces the non-compact Minkowski spacetime $\mathbb{R}^4$ with the compact Euclidean spacetime $\mathbb{S}^4$. In our notation the Latin spacetime indices $m, n, \ldots$ are used regardless of the type of spacetime, which is specified elsewhere. To continue from Minkowski spacetime to Euclidean spacetime, the temporal co-ordinate x^0 is Wick rotated to the imaginary axis, with the result that:

$$x^0 \to -ix^0, \ \ \partial_0 \to i\partial_0, \tag{A.3}$$

leaving the spatial co-ordinates x^i, $i = 1,2,3$ unaffected. The gauge field potential v_m in Minkowski spacetime requires the component transformations:

$$v_0 \to iv_0, \ \ v_i \to v_i, \tag{A.4}$$

for continuation to Euclidean spacetime. Similar transformation can be used so that the definitions of the gauge covariant derivative and gauge field strength in Minkowski spacetime remain intact in Euclidean spacetime. The standard Minkowski spacetime gamma matrices can also be continued to Euclidean spacetime, enabling one to continue fermionic terms which appear in Minkowski spacetime actions to Euclidean spacetime actions, through the following transformations:

$$\gamma_0 \rightarrow \gamma_0, \quad \gamma_i \rightarrow i\gamma_i. \tag{A.5}$$

The reality of the Minkowski spacetime is also altered by analytic continuation to Euclidean spacetime. Given a Minkowski spacetime Yang–Mills action S_M, and the same action in Euclidean spacetime S_E, one has:

$$S_M = iS_E. \tag{A.6}$$

This modification of the reality of S_E does not affect the physical results of the path integration, which remains convergent.

Weyl Spinors

The conventions for generic Weyl spinors in this thesis are as follows. A left-handed Weyl spinor transforms in the $(\frac{1}{2},0)$ representation of the Lorentz group, $SO(3,1)$, and is denoted ψ_α. A right-handed Weyl spinor is the spinor conjugate of ψ_α, which transforms in the $(0,\frac{1}{2})$ representation of $SO(3,1)$ and is denoted by $\bar{\psi}^{\dot{\alpha}}$. Weyl spinor indices can be raised or lowered using the antisymmetric tensors $\epsilon^{\alpha\beta}$ and $\epsilon_{\alpha\beta}$, explicitly given by:

$$\epsilon^{\alpha\beta} = \epsilon^{\dot{\alpha}\dot{\beta}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \epsilon_{\alpha\beta} = \epsilon_{\dot{\alpha}\dot{\beta}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{A.7}$$

which also obey $\epsilon_{\alpha\beta}\epsilon^{\beta\gamma} = \delta^\gamma_\alpha$. Two Weyl spinors are contracted using the rules:

$$\chi\psi = \chi^\alpha\psi_\alpha, \quad \bar{\chi}\bar{\psi} = \bar{\chi}_{\dot{\alpha}}\bar{\psi}^{\dot{\alpha}}. \tag{A.8}$$

The sigma matrices in Euclidean spacetime are defined in quaternionic notation as:

$$\sigma^n_{\alpha\dot{\alpha}} = (-1_{[2]\times[2]}, \tau^c), \quad \bar{\sigma}^{n\dot{\alpha}\alpha} = (-1_{[2]\times[2]}, -\tau^c), \tag{A.9}$$

where τ^c, $c = 1, 2, 3$ are the three standard Pauli matrices in four dimensional Euclidean spacetime, given by:

$$\tau^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{A.10}$$

The antisymmetric tensors in Eqs. (A.7) can be used to relate the sigma matrices, via:

$$\sigma^n_{\alpha\dot{\alpha}} = \epsilon_{\alpha\beta}\epsilon_{\dot{\alpha}\dot{\beta}}\bar{\sigma}^{n\dot{\beta}\beta}. \tag{A.11}$$

Dirac Spinors

A Dirac, or two-component, spinor, can be composed of two Weyl spinors. We denote a generic Dirac spinor by Ψ, and its two Weyl spinor components as ψ_α and $\bar{\chi}^{\dot{\alpha}}$, where:

$$\Psi = \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix}. \tag{A.12}$$

The gamma matrices γ^n associated with the covariant derivatives of spinors can be written in terms of sigma matrices as:

$$\gamma^n = \begin{pmatrix} 0 & \sigma^n \\ \bar{\sigma}^n & 0 \end{pmatrix}. \tag{A.13}$$

Grassmann Spinors

In the study of $\mathcal{N} = 1$ supersymmetry, it is convenient to use constant Grassmann spinors θ^α and $\bar{\theta}_{\dot{\alpha}}$, where $\alpha, \beta, \dot{\alpha}, \dot{\beta} = 1, 2$ are Weyl indices, to write the $\mathcal{N} = 1$ supersymmetry algebra as a Lie algebra. These spinors are said to be Grassmann-valued if they obey the following Grassmann algebra:

$$\{\theta^\alpha, \theta^\beta\} = \{\bar{\theta}_{\dot{\alpha}}, \bar{\theta}_{\dot{\beta}}\} = \{\theta^\alpha, \bar{\theta}_{\dot{\beta}}\} = 0. \tag{A.14}$$

in which $\{A, B\} = AB + BA$ is the standard anticommutator bracket. The spinors θ^α and $\bar{\theta}_{\dot{\alpha}}$ obey the following algebra, which involve contraction over the Weyl indices:

$$\theta^\alpha\theta^\beta = -\tfrac{1}{2}\epsilon^{\alpha\beta}\theta\theta, \quad \theta_\alpha\theta_\beta = \tfrac{1}{2}\epsilon^{\alpha\beta}\theta\theta, \tag{A.15}$$

$$\bar{\theta}^{\dot{\alpha}}\bar{\theta}^{\dot{\beta}} = \tfrac{1}{2}\epsilon^{\dot{\alpha}\dot{\beta}}\bar{\theta}\bar{\theta}, \quad \bar{\theta}_{\dot{\alpha}}\bar{\theta}_{\dot{\beta}} = -\tfrac{1}{2}\epsilon_{\dot{\alpha}\dot{\beta}}\bar{\theta}\bar{\theta}, \tag{A.16}$$

and in which the antisymmetric tensors given in Eq. (A.7) have been used. There are many other formulae which the constant Grassmann spinors θ^α, $\bar{\theta}_{\dot{\alpha}}$ obey. These include the rules for the differentiation and integration of Grassmann-valued variables; for details of these we refer the reader to the reviews [123, 128] and the book [122].

Appendix B

Properties of Quaternions

A quaternion can be defined in a number of ways. A quaternion can be realized as a set of 2×2 complex matrices with real entries or as an ordered quadruple of real numbers. Quaternions form one of the real division algebras, often denoted by $\mathbf{Q}$ [284]; the set of all quaternions is usually denoted by $\mathbb{H}$. Quaternions are also referred to as hyper-complex numbers, but complex numbers do not generalize to quaternions in all respects. In the following we denote generic quaternions as p, q and r. A quaternion q can be written in component form as:

$$q = q_0 + q_1 i + q_2 j + q_3 k, \tag{B.1}$$

where $\{i, j, k\}$ are basis quaternions and $\{q_0, q_1, q_2, q_3\} \in \mathbb{R}$ are real numbers. The actual basis of the quaternion is $\{1, i, j, k\}$, which leads to the representation of the set of quaternions as a vector space of dimension four over the real field $\mathbb{R}$. The basis quaternions $\{i, j, k\}$ are numbers which satisfy the following properties:

$$i^2 = j^2 = k^2 = ijk = -1, \tag{B.2}$$

$$ij = -ji = k, \tag{B.3}$$

$$jk = -kj = i, \tag{B.4}$$

$$ki = -ik = j. \tag{B.5}$$

The real part of the quaternion q is given by:

$$\mathrm{Re}(q) = q_0, \tag{B.6}$$

whilst q has three 'imaginary' parts, given by:

$$\mathrm{Im}(q) = \{q_1, q_2, q_3\}, \tag{B.7}$$

however, the use of the term 'imaginary' is used in analogy with complex numbers and clearly is not congruent to the imaginary part of a complex number.

Two different quaternions q and p are equal if and only if their real and imaginary parts are equal; thus:

$$p = q \Rightarrow q_0 = p_0, \quad \{q_1, q_2, q_3\} = \{p_1, p_2, p_3\}. \tag{B.8}$$

The negative of a quaternion is implemented by negating all of its components:

$$-q = -q_0 - q_1 i - q_2 j - q_3 k. \tag{B.9}$$

The conjugate of a quaternion is not the same as its negative. The quaternionic conjugate of a quaternion is given by the quaternion with all of its 'imaginary' parts negated. We denote the quaternionic conjugate of q by $\bar{q}$, where:

$$\bar{q} = q_0 - q_1 i - q_2 j - q_3 k. \tag{B.10}$$

This is essentially equivalent to the complex conjugation of each of the basis quaternions $\{i, j, k\}$. Quaternions obey the ordinary laws of linear algebra except for multiplication. The multiplication of quaternions makes use of the distributive law in Eq. (B.2). We consider quaternion multiplication after describing quaternion addition and subtraction. Addition of quaternions is effected by the addition of the separate components of the quaternions, retaining the ordering of the components. The addition of the quaternions q and p is given by:

$$p + q = (p_0 + q_0) + (p_1 + q_1)i + (p_2 + q_2)j + (p_3 + q_3)k. \tag{B.11}$$

Quaternion addition is associative:

$$p + (q + r) = (p + q) + r, \tag{B.12}$$

$$p + q = q + p. \tag{B.13}$$

Subtraction of quaternions is defined by:

$$p - q = p + (-q). \tag{B.14}$$

The magnitude or absolute value of a quaternion p is a real number defined as the non-negative square root of the sum of the squares of the components of p. The magnitude of a quaternion p is denoted by $|p|$, where:

$$|p| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}. \tag{B.15}$$

A quaternion p can only have zero magnitude, $|p|^2 = 0$, if and only if it is identical to the quaternion with all zero components; that is, $|p|^2 = 0 \Rightarrow p \equiv 0 + 0i + 0j + 0k$. The quaternion p is a unit quaternion if each of its components are divided by its magnitude. The direction of the unit quaternion is given by the orientation of the quaternion whose components are divided by another quaternion. The unit quaternion in the direction of p is defined by:

$$\frac{p}{|p|} = \frac{p_0}{|p|} + \frac{p_1}{|p|}i + \frac{p_2}{|p|}j + \frac{p_3}{|p|}k. \tag{B.16}$$

Quaternion multiplication is a non-trivial operation. Given two generic quaternions p and q, p and q multiply according to:

$$\begin{aligned} pq =& (p_0q_0 - p_1q_1 - p_2q_2 - p_3q_3) + (p_0q_1 + p_1q_0 + p_2q_3 - p_3q_2)i+ \\ & (p_0q_2 + p_2q_0 + p_3q_1 - p_1q_3)j + (p_0q_3 + p_3q_0 - p_2q_1 + p_1q_2)k. \end{aligned} \tag{B.17}$$

Significantly, quaternion multiplication is noncommutative: in general, $pq \neq qp$. This is due to the distributive law in Eq. (B.2). Quaternions obey the following multiplication laws:

$$p(qr) = (pq)r, \tag{B.18}$$

$$p(q + r) = pq + pr, \tag{B.19}$$

$$(p + q)r = pr + qr, \tag{B.20}$$

$$\overline{pq} = \bar{q} \cdot \bar{p} \tag{B.21}$$

$$|pq| = |p||q|. \tag{B.22}$$

Furthermore, if $pq = 0$, then $p = 0$ or $q = 0$ or $p = q = 0$. We also note that the square of the magnitude of p may be obtained by the product of a quaternion and its conjugate:

$$|p|^2 = p\bar{p} = \bar{p}p. \tag{B.23}$$

Two quaternions p and q are said to be proportional in their imaginaries if there exist five real numbers $\{a, b, c, d, e\}$ such that:

$$p = p_0 + adi + bdj + cdk, \tag{B.24}$$

$$q = q_0 + aei + bej + cek, \tag{B.25}$$

in which the values of the real parts of p and q are irrelevant. Quaternions p and q commute when multiplied together if and only if they are proportional in their imaginaries, as in Eqs. (B.24,B.25). Further results regarding quaternion multiplication include the following. If pq is equal to a real number, then p and q commute. If p commutes with q, then p also commutes with pq. If $pq = qp$ and $pr = rp$, then $p(q + r) = (q + r)p$. If $F(p)$ is a polynomial in the quaternion p, then $F(p)p = pF(p)$. If $pq = qp$ for all quaternions q, then p is real. Given a quaternion p, one may specify the components of p separately by applying the tetranomial formulae:

$$p_0 = \tfrac{1}{4}[p - ipi - jpj - kpk], \tag{B.26}$$

$$p_1 = -\tfrac{1}{4}[p - ipi + jpj + kpk]i, \tag{B.27}$$

$$p_0 = -\tfrac{1}{4}[p + ipi - jpj + kpk]j, \tag{B.28}$$

$$p_0 = -\tfrac{1}{4}[p + ipi + jpj - kpk]k. \tag{B.29}$$

The result is the tetrasection of p, which gives the components of p as functions of p. The complex numbers have no analogue for these formulae. Multiplication of positive powers of quaternions are associative:

$$p^n p^m = p^m p^n = p^{m+n}, \tag{B.30}$$

$$(p^n)^m = p^{nm}. \tag{B.31}$$

Division of one quaternion by another quaternion is not defined in general. Division of a quaternion p by a real number follows the same rule as multiplication of p by a real number. For a real number $W \in \mathbb{R}$, one has:

$$Wp = Wp_0 + Wp_1 i + Wp_2 j + Wp_3 k, \quad \frac{1}{W}p = \frac{1}{W}p_0 + \frac{1}{W}p_1 i + \frac{1}{W}p_2 j + \frac{1}{W}p_3 k. \tag{B.32}$$

It follows that the unit quaternion for non-zero p can be written as $p/|p|$. The quaternions have a reciprocal or multiplicative inverse, which is defined for a non-zero quaternion p

as $p^{-1} = 1/p$, where:

$$\frac{1}{p} \cdot p = p \cdot \frac{1}{p} = 1, \quad \frac{1}{p} = \frac{\bar{p}}{|p|^2}, \quad \frac{1}{pq} = \frac{1}{p} \cdot \frac{1}{q}. \tag{B.33}$$

Polynomials of quaternions behave in a markedly different way to complex polynomials. For example, given a constant quaternion a and a variable quaternion q, a first degree monomial of the form:

$$aq = qb, \tag{B.34}$$

cannot be satisfied by a constant quaternion b. Hence, although a is a constant, b must vary. This is because b cannot assume two values simultaneously. A generic monomial term for quaternions may be taken as aqb, with $\{a, b, q\}$ all quaternions. The study of polynomials of quaternions is not well established. For instance, given a polynomial of quaternions $aqb + cqd$, there are in general no constant quaternions r and s such that

$$aqb + cqd = rqs. \tag{B.35}$$

Although there are no quaternions r and s which solve Eq. (B.35), one can write Eq. (B.35) in terms of quaternion components, with the result that Eq. (B.35) becomes four real linear equations. The theory of linear algebra can then be applied to these equations, and r and s can be solved for. However, the quaternions r and s cannot then be reconstructed from the resulting component form solution, as constant quaternions r and s do not exist which solve Eq. (B.35).

Problems associated with quaternion polynomials also extend to quaternion polynomials of higher degree. Given a variable quaternion q and a constant quaternion a, the second degree equation $q^2 = a$ can only be solved in general by considering the separate cases in which a is either zero, a positive real number, a negative real number, or a complex number, and using real analysis to solve each case in turn.

Bibliography

[1] A. A. Belavin, A. M. Polyakov, A. S. Shvarts and Y. S. Tyupkin, Phys. Lett. B **59** (1975) 85.

[2] A. M. Polyakov, Phys. Lett. B **59** (1975) 82.

[3] R. S. Ward, Phys. Lett. A **61** (1977) 81.

[4] M. F. Atiyah and R. S. Ward, Commun. Math. Phys. **55** (1977) 117.

[5] M. F. Atiyah, N. J. Hitchin and I. M. Singer, Proc. Nat. Acad. Sci. **74** (1977) 2662.

[6] M. F. Atiyah and J. D. Jones, Commun. Math. Phys. **61** (1978) 97.

[7] M. F. Atiyah, N. J. Hitchin and I. M. Singer, Proc. Roy. Soc. Lond. A **362** (1978) 425.

[8] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld and Y. I. Manin, Phys. Lett. A **65** (1978) 185.

[9] V. G. Drinfeld and Y. I. Manin, Commun. Math. Phys. **63** (1978) 177.

[10] R. Hartshorne, Commun. Math. Phys. **59** (1978) 1.

[11] R. Bott, Bull. Soc. Math. Fr. **84** (1956) 251.

[12] M. F. Atiyah and I. M. Singer, Annals Math. **87** (1968) 484,
M. F. Atiyah and G. B. Segal, Annals Math. **87** (1968) 531,
M. F. Atiyah and I. M. Singer, Annals Math. **87** (1968) 546,
M. F. Atiyah and I. M. Singer, Annals Math. **93** (1971) 119,

M. F. Atiyah and I. M. Singer, Annals Math. **93** (1971) 139.
M. Atiyah, R. Bott and V. K. Patodi, Invent. Math. **19** (1973) 279.

[13] M. F. Atiyah, "Geometry Of Yang–Mills Fields," *Pisa, Italy: Sc. Norm. Sup., 1979 (Lezioni Fermiane).*

[14] M. Atiyah, "Michael Atiyah Collected Works. Vol. 5: Gauge Theories," *Oxford, U.K.: Clarendon (1988).*

[15] A. M. Polyakov, Nucl. Phys. B **120** (1977) 429.

[16] R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37** (1976) 172.

[17] C. G. Callan, R. F. Dashen and D. J. Gross, Phys. Lett. B **63** (1976) 334,
C. G. Callan, R. F. Dashen and D. J. Gross, Phys. Lett. B **66** (1977) 375,
C. G. Callan, R. F. Dashen and D. J. Gross, Phys. Rev. D **17** (1978) 2717,
C. G. Callan, R. F. Dashen and D. J. Gross, Phys. Rev. D **20** (1979) 3279,
C. G. Callan, R. F. Dashen and D. J. Gross, Phys. Rev. Lett. **44** (1980) 435.

[18] G. 't Hooft, unpublished.

[19] R. Jackiw and C. Rebbi, Phys. Rev. D **14** (1976) 517,
R. Jackiw, C. Nohl and C. Rebbi, Phys. Rev. D **15** (1977) 1642.

[20] E. Witten, Phys. Rev. Lett. **38** (1977) 121.

[21] C. K. Peng, Sci. Sin. **20** (1977) 345.

[22] N. H. Christ, E. J. Weinberg and N. K. Stanton, Phys. Rev. D **18** (1978) 2013.

[23] V. E. Korepin and S. L. Shatashvili, Sov. Phys. Dokl. **28** (1983) 1018.

[24] V. I. Inozemtsev, "On Classical Yang–Mills Fields With Symplectic Gauge Groups And Topological Charge K = 4," preprint JINR-P2-82-491.

[25] F. Wilczek, Phys. Lett. B **65** (1976) 160.

[26] C. W. Bernard, N. H. Christ, A. H. Guth and E. J. Weinberg, Phys. Rev. D **16** (1977) 2967.

[27] H. Arfaei, Phys. Lett. B **76** (1978) 292.

[28] A. S. Schwarz, Lett. Math. Phys. **2** (1978) 201.

[29] A. A. Belavin, V. A. Fateev, A. S. Schwarz and Y. S. Tyupkin, Phys. Lett. B **83** (1979) 317.

[30] E. Corrigan and D. B. Fairlie, Phys. Lett. B **67** (1977) 69.

[31] E. Corrigan, D. B. Fairlie, R. G. Yates and P. Goddard, Phys. Lett. B **72** (1978) 354.

[32] E. Corrigan, D. B. Fairlie, R. G. Yates and P. Goddard, Commun. Math. Phys. **58** (1978) 223.

[33] E. Corrigan, D. B. Fairlie, S. Templeton and P. Goddard, Nucl. Phys. B **140** (1978) 31.

[34] E. Corrigan, P. Goddard and S. Templeton, Nucl. Phys. B **151** (1979) 93.

[35] V. G. Drinfeld and Y. I. Manin, Yad. Fiz. **29** (1979) 1646.

[36] N. B. Pomeroy, Phys. Lett. B **547** (2002) 85 [arXiv:hep-th/0203184].

[37] G. 't Hooft, Phys. Rev. D **14** (1976) 3432 [Erratum-ibid. D **18** (1978) 2199].

[38] L. S. Brown, R. D. Carlitz and C. K. Lee, Phys. Rev. D **16** (1977) 417.
L. S. Brown and C. K. Lee, Phys. Rev. D **18** (1978) 2180.

[39] C. W. Bernard and E. J. Weinberg, Phys. Rev. D **15** (1977) 3656,
C. W. Bernard, Phys. Rev. D **18** (1978) 2026,
C. W. Bernard, "Physical Effects Of Instantons," UCLA/79/TEP/12 *talks given at Canadian Math. Society Summer Research Inst., Geometrical and Topological Methods in Gauge Theories, Sep 1979.*

[40] C. W. Bernard, Phys. Rev. D **19** (1979) 3013.

[41] H. Osborn, Nucl. Phys. B **140** (1978) 45.

[42] H. Osborn, Nucl. Phys. B **159** (1979) 497.

[43] I. Jack, Nucl. Phys. B **174** (1980) 526.

[44] P. Goddard, P. Mansfield and H. Osborn, Phys. Lett. B **98** (1981) 59.

[45] H. Osborn and G. P. Moody, Nucl. Phys. B **173** (1980) 422.

[46] P. Mansfield, Nucl. Phys. B **186** (1981) 287.

[47] H. Osborn, Annals Phys. **135** (1981) 373.

[48] I. Jack and H. Osborn, Nucl. Phys. B **207** (1982) 474,
I. Jack and H. Osborn, Nucl. Phys. B **249** (1985) 472.

[49] D. I. Olive, S. Sciuto and R. J. Crewther, Riv. Nuovo Cim. **2N8** (1979) 1.

[50] E. Corrigan, Phys. Rept. **49** (1979) 95.

[51] V. G. Drinfeld, I. M. Krichever, Y. I. Manin and S. P. Novikov, "Methods Of Algebraic Geometry In Contemporary Mathematical Physics," *In Novikov, S. P. ed. : Mathematical Physics Reviews, Vol. 1, 1-54.*

[52] H. Osborn, Commun. Math. Phys. **86** (1982) 195.

[53] E. Corrigan and P. Goddard, Annals Phys. **154** (1984) 253.

[54] C. W. Bernard, A. H. Guth and E. J. Weinberg, Phys. Rev. D **17** (1978) 1053.

[55] A. Patrascioiu, Phys. Rev. D **17** (1978) 2764.

[56] D. R. Stump, Phys. Lett. B **76** (1978) 307.

[57] D. R. Stump, Phys. Rev. D **17** (1978) 2126.

[58] W. Pauli and F. Villars, Rev. Mod. Phys. **21** (1949) 434.

[59] E. Corrigan, P. Goddard, H. Osborn and S. Templeton, Nucl. Phys. B **159** (1979) 469.

[60] S. Chadha, P. Di Vecchia, A. D'Adda and F. Nicodemi, Phys. Lett. B **72** (1977) 103,
S. Chadha, A. D'Adda, P. Di Vecchia and F. Nicodemi, Phys. Lett. B **67** (1977) 103.

[61] A. S. Schwarz, Commun. Math. Phys. **64** (1979) 233.

[62] K. Huleihil, Phys. Lett. B **138** (1984) 287.

[63] S. R. Coleman, "The Uses Of Instantons," in "Aspects of Symmetry: selected Erice lectures of Sidney Coleman," *Cambridge, U.K.: Cambridge University Press, 1985.*

[64] A. I. Vainshtein, V. I. Zakharov, V. A. Novikov and M. A. Shifman, Sov. Phys. Usp. **24** (1982) 195 [Usp. Fiz. Nauk **136** (1982) 553].

[65] R. Flume, Phys. Lett. B **76** (1978) 593.

[66] D. H. Tchrakian, J. Math. Phys. **21** (1980) 166.

[67] N. J. Hitchin, A. Karlhede, U. Lindstrom and M. Rocek, Commun. Math. Phys. **108** (1987) 535.

[68] A. Maciocia, Commun. Math. Phys. **135** (1991) 467.

[69] E. Brezin and J. L. Gervais, Phys. Rept. **49** (1979) 91.

[70] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **165** (1980) 45.

[71] T. Schafer and E. V. Shuryak, Rev. Mod. Phys. **70** (1998) 323 [arXiv:hep-ph/9610451],
E. V. Shuryak and I. Zahed, Phys. Rev. D **62** (2000) 085014 [arXiv:hep-ph/0005152],
E. V. Shuryak and T. Schafer, Nucl. Phys. Proc. Suppl. **53** (1997) 472.

[72] D. Amati, Nucl. Phys. B **279** (1987) 295.

[73] M. Hutter, [arXiv:hep-ph/0107098].

[74] B. Metsch, Nucl. Phys. A **675** (2000) 161C,
T. G. Kovacs, Nucl. Phys. Proc. Suppl. **83** (2000) 500 [arXiv:hep-lat/9908055],
W. H. Blask, U. Bohn, M. G. Huber, B. C. Metsch and H. R. Petry, Z. Phys. A **337**

(1990) 327,

D. Horn and S. Yankielowicz, Phys. Lett. B **76** (1978) 343.

[75] A. Actor, Rev. Mod. Phys. **51** (1979) 461.

[76] R. Jackiw, Rev. Mod. Phys. **52** (1980) 661.

[77] E. Kovacs and S. y. Lo, Phys. Rev. D **19** (1979) 2966.

[78] G. H. Derrick, J. Math. Phys. **5** (1964) 1252.
R. Hobart, Proc. Royal Soc. London **A82** (1963) 201.

[79] I. Affleck, Nucl. Phys. B **191**, 429 (1981).

[80] M. Nielsen and N. K. Nielsen, [arXiv:hep-th/9912006].

[81] I. Affleck, Nucl. Phys. B **162**, 461 (1980),
I. Affleck, Nucl. Phys. B **171**, 420 (1980).

[82] I. Affleck, Phys. Lett. B **92**, 149 (1980).

[83] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Phys. Lett. B **76** (1978) 471.

[84] A. A. Belavin and V. E. Zakharov, Pisma Zh. Eksp. Teor. Fiz. **25** (1977) 603,
A. A. Belavin and V. E. Zakharov, Phys. Lett. B **73** (1978) 53.

[85] A. E. Arinshtein, Yad. Fiz. **29** (1979) 249.

[86] A. A. Belavin, Sov. Sci. Rev. A **1** (1979) 1.

[87] K. Takasaki, Commun. Math. Phys. **94** (1984) 35.

[88] C. Nash, Phys. Lett. B **74**, 73 (1978).

[89] D. E. Lerner and P. D. Sommers, "Complex Manifold Techniques In Theoretical Physics. Proceedings, Workshop, Lawrence, U.S.A., July 10-15, 1978," *San Francisco, U.S.A.: Pitman (1979) (Research Notes In Mathematics, 32).*

[90] R. Rajaraman, "Solitons And Instantons. An Introduction To Solitons And Instantons In Quantum Field Theory," *Amsterdam, Netherlands: North-Holland (1982).*

[91] Y. I. Manin, "Gauge Field Theory And Complex Geometry," *Berlin, Germany: Springer (1988) (Grundlehren Der Mathematischen Wissenschaften, 289).*

[92] R. S. Ward and R. O. Wells, "Twistor Geometry And Field Theory," *Cambridge, U.K.: Cambridge University Press, 1990.*

[93] J. Madore, J. L. Richard and R. Stora, Phys. Rept. **49** (1979) 113.

[94] J. Madore, Phys. Rept. **75** (1981) 125.

[95] J. Spruck, D. H. Tchrakian and Y. Yang, Commun. Math. Phys. **188** (1997) 737.

[96] M. A. Singer and P. M. Sutcliffe, Nonlinearity **12** (1999) 987 [arXiv:hep-th/9901075].

[97] A. V. Belitsky, S. Vandoren and P. van Nieuwenhuizen, Class. Quant. Grav. **17** (2000) 3521 [arXiv:hep-th/0004186].

[98] S. Vandoren, [arXiv:hep-th/0009150].

[99] A. V. Belitsky, S. Vandoren and P. van Nieuwenhuizen, Phys. Lett. B **477** (2000) 335 [arXiv:hep-th/0001010].

[100] J. Isenberg, P. B. Yasskin and P. S. Green, Phys. Lett. B **78** (1978) 462.

[101] E. Witten, Phys. Lett. B **77** (1978) 394.

[102] C. H. Oh and R. Teh, Phys. Rev. D **21** (1980) 531.

[103] A. L. Koshkarov, Theor. Math. Phys. **105** (1996) 1534 [Teor. Mat. Fiz. **105N3** (1995) 423].

[104] W. Nahm, Phys. Lett. B **90** (1980) 413.

[105] H. Panagopoulos, Phys. Rev. D **28** (1983) 380.

[106] E. Corrigan, P. R. Wainwright and S. M. Wilson, Commun. Math. Phys. **98** (1985) 259.

[107] P. M. Sutcliffe, Int. J. Mod. Phys. A **12** (1997) 4663 [arXiv:hep-th/9707009].

[108] A. Galperin and E. Sokatchev, Nucl. Phys. B **452**, 431 (1995) [arXiv:hep-th/9412032].

[109] S. R. Coleman and J. Mandula, Phys. Rev. **159** (1967) 1251.

[110] D. V. Volkov and V. P. Akulov, Phys. Lett. B **46** (1973) 109.

[111] Y. A. Golfand and E. P. Likhtman, JETP Lett. **13** (1971) 323 [Pisma Zh. Eksp. Teor. Fiz. **13** (1971) 452].

[112] P. Fayet, Nucl. Phys. B **113** (1976) 135.

[113] J. Wess and B. Zumino, Phys. Lett. B **49** (1974) 52,
S. Ferrara, J. Wess and B. Zumino, Phys. Lett. B **51**, 239 (1974),
J. Wess and B. Zumino, Nucl. Phys. B **70**, 39 (1974).

[114] A. Salam and J. Strathdee, Nucl. Phys. B **76** (1974) 477,
A. Salam and J. Strathdee, Nucl. Phys. B **80** (1974) 499,
A. Salam and J. Strathdee, Phys. Rev. D **11** (1975) 1521,
A. Salam and J. Strathdee, Phys. Lett. B **51** (1974) 353.

[115] A. Salam and J. Strathdee, Fortsch. Phys. **26** (1978) 57.

[116] R. Haag, J. T. Lopuszanski and M. Sohnius, Nucl. Phys. B **88** (1975) 257.

[117] L. Brink, J. H. Schwarz and J. Scherk, Nucl. Phys. B **121** (1977) 77.

[118] R. Grimm, M. Sohnius and J. Wess, Nucl. Phys. B **133** (1978) 275.
R. Grimm, J. Wess and B. Zumino, Nucl. Phys. B **152**, 255 (1979).

[119] M. F. Sohnius, Nucl. Phys. B **136** (1978) 461,
M. F. Sohnius, Nucl. Phys. B **138** (1978) 109.

[120] P. Fayet and S. Ferrara, Phys. Rept. **32** (1977) 249.
S. Ferrara, ed., "Supersymmetry," *Amsterdam, Holland: North-Holland, (1987).*

[121] M. F. Sohnius, Phys. Rept. **128** (1985) 39.

[122] J. Bagger and J. Wess, "Supersymmetry And Supergravity," *Princeton: Princeton University Press (1982). Princeton Series in Physics.*

[123] J. D. Lykken, [arXiv:hep-th/9612114].

[124] M. Dine, [arXiv:hep-ph/9612389].

[125] S. P. Martin, [arXiv:hep-ph/9709356].

[126] P. C. West, "Introduction To Supersymmetry And Supergravity," *Singapore, Singapore: World Scientific (1990).*

[127] S. J. Gates, M. T. Grisaru, M. Rocek and W. Siegel, "Superspace, Or One Thousand And One Lessons In Supersymmetry," Front. Phys. **58** (1983) 1 [arXiv:hep-th/0108200].

[128] F. A. Berezin, Sov. J. Phys. **29** (1979) [Yad. Fiz. **29** (1979) 1670].

[129] P. Van Nieuwenhuizen, Phys. Rept. **68** (1981) 189.

[130] I. Affleck, M. Dine and N. Seiberg, Phys. Rev. Lett. **51**, 1026 (1983).

[131] E. Witten, Nucl. Phys. B **188**, 513 (1981).

[132] Y. Shadmi and Y. Shirman, Rev. Mod. Phys. **72** (2000) 25 [arXiv:hep-th/9907225].

[133] A. Galperin, E. Ivanov, S. Kalitsyn, V. Ogievetsky and E. Sokatchev, Phys. Lett. B **151** (1985) 215,
A. Galperin, E. Ivanov, S. Kalitsyn, V. Ogievetsky and E. Sokatchev, Class. Quant. Grav. **2** (1985) 155,
A. S. Galperin, E. A. Ivanov and V. I. Ogievetsky, Sov. J. Nucl. Phys. **46** (1987) 543 [Yad. Fiz. **46** (1987) 948].

[134] A. S. Galperin, E. A. Ivanov, V. I. Ogievetsky and E. S. Sokatchev, "Harmonic Superspace," *Cambridge, U. K. Cambridge University Press (2001).*
W. Siegel, Class. Quant. Grav. **2** (1985) 439.

[135] M. T. Grisaru, W. Siegel and M. Rocek, Nucl. Phys. B **159** (1979) 429.

[136] V. A. Novikov, M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **229** (1983) 381.

[137] V. A. Novikov, M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Phys. Lett. B **166** (1986) 329.

[138] M. A. Shifman, Prog. Part. Nucl. Phys. **39** (1997) 1 [arXiv:hep-th/9704114], M. A. Shifman, Int. J. Mod. Phys. A **14** (1999) 5017 [arXiv:hep-th/9906049], M. A. Shifman, ed., "Instantons In Gauge Theories," *Singapore ; River Edge, NJ: World Scientific, (1994).*

[139] M. A. Shifman and A. I. Vainshtein, [arXiv:hep-th/9902018], M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Sov. Phys. Usp. **28** (1985) 709 [Usp. Fiz. Nauk **146** (1985) 683].

[140] J. Iliopoulos, Nucl. Phys. Proc. Suppl. **101** (2001) 99 [arXiv:hep-th/0103210].

[141] N. Seiberg, Phys. Lett. B **318** (1993) 469 [arXiv:hep-ph/9309335].

[142] S. F. Cordes, Nucl. Phys. B **273** (1986) 629.

[143] T. Ohrndorf, Nucl. Phys. B **273** (1986) 165.

[144] N. M. Davies, T. J. Hollowood and V. V. Khoze, [arXiv:hep-th/0006011], N. M. Davies and V. V. Khoze, JHEP **0001**, 015 (2000) [arXiv:hep-th/9911112], N. M. Davies, T. J. Hollowood, V. V. Khoze and M. P. Mattis, Nucl. Phys. B **559**, 123 (1999) [arXiv:hep-th/9905015].

[145] N. Seiberg, Phys. Rev. D **49** (1994) 6857 [arXiv:hep-th/9402044].

[146] K. A. Intriligator, R. G. Leigh and N. Seiberg, Phys. Rev. D **50** (1994) 1092 [arXiv:hep-th/9403198].

[147] K. A. Intriligator and N. Seiberg, Nucl. Phys. B **431** (1994) 551 [arXiv:hep-th/9408155].

[148] P. A. Dirac, Proc. Roy. Soc. Lond. A **133** (1931) 60.

[149] A. M. Polyakov, JETP Lett. **20** (1974) 194 [Pisma Zh. Eksp. Teor. Fiz. **20** (1974) 430],
G. 't Hooft, Nucl. Phys. B **79** (1974) 276.

[150] E. B. Bogomolny, Sov. J. Nucl. Phys. **24** (1976) 449 [Yad. Fiz. **24** (1976) 861].
M. K. Prasad and C. M. Sommerfield, Phys. Rev. Lett. **35** (1975) 760.
E. Corrigan, D. I. Olive, D. B. Fairlie and J. Nuyts, Nucl. Phys. B **106**, 475 (1976).

[151] P. Goddard, J. Nuyts and D. I. Olive, Nucl. Phys. B **125** (1977) 1.

[152] C. Montonen and D. I. Olive, Phys. Lett. B **72** (1977) 117.

[153] D. I. Olive, Phys. Rept. **49** (1979) 165.

[154] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32** (1974) 438.
H. Georgi and S. L. Glashow, Phys. Rev. D **6** (1972) 2977.

[155] S. Ferrara, D. Lust, A. D. Shapere and S. Theisen, Phys. Lett. B **225** (1989) 363.

[156] A. Sen, Int. J. Mod. Phys. A **9** (1994) 3707 [arXiv:hep-th/9402002],
A. Sen, Phys. Lett. B **329** (1994) 217 [arXiv:hep-th/9402032].

[157] N. Dorey, C. Fraser, T. J. Hollowood and M. A. Kneipp, Phys. Lett. B **383**, 422 (1996) [arXiv:hep-th/9605069].
N. Dorey, C. Fraser, T. J. Hollowood and M. A. Kneipp, [arXiv:hep-th/9512116].

[158] H. Osborn, Phys. Lett. B **83** (1979) 321.

[159] J. M. Figueroa-O'Farrill, "Electric-magnetic Duality for Children," `http://www.maths.ed.ac.uk/ jmf/`.

[160] A. D'Adda, R. Horsley and P. Di Vecchia, Phys. Lett. B **76** (1978) 298.

[161] H. Osborn, Phys. Lett. B **115** (1982) 226.

[162] R. Barbieri, S. Ferrara, L. Maiani, F. Palumbo and C. A. Savoy, Phys. Lett. B **115** (1982) 212.

[163] P. S. Howe, K. S. Stelle and P. C. West, Phys. Lett. B **124** (1983) 55.

[164] P. Di Vecchia, R. Musto, F. Nicodemi and R. Pettorino, Nucl. Phys. B **252** (1985) 635.

[165] P. S. Howe, A. J. Parkes and P. C. West, Phys. Lett. B **150** (1985) 149.

[166] A. W. Fisher, Nucl. Phys. B **229** (1983) 142.

[167] E. Witten and D. I. Olive, Phys. Lett. B **78** (1978) 97.

[168] E. Witten, Phys. Lett. B **86** (1979) 283.

[169] N. Seiberg, Phys. Lett. B **206** (1988) 75.

[170] N. Seiberg and E. Witten, Nucl. Phys. B **426** (1994) 19 [Erratum-ibid. B **430** (1994) 485] [arXiv:hep-th/9407087].

[171] N. Seiberg and E. Witten, Nucl. Phys. B **431** (1994) 484 [arXiv:hep-th/9408099].

[172] E. Witten, "Field Theory," *lectures given at the IAS 'Quantum Field Theory for Mathematicians' workshop, Institute of Advanced Studies, Princeton, U.S.A., 1996.* `http://www.math.ias.edu/QFT/spring/index.html`,
"Quantum Fields and Strings: A Course For Mathematicians," P. Deligne, P. Etingof, D.S. Freed, L. Jeffrey, D. Kazhdan, J. Morgan, D.R. Morrison and E. Witten, eds., *American Mathematical Society, Providence, U.S.A., 1999*.

[173] A. Klemm, W. Lerche, S. Yankielowicz and S. Theisen, Phys. Lett. B **344**, 169 (1995) [arXiv:hep-th/9411048],
A. Klemm, W. Lerche, S. Yankielowicz and S. Theisen, [arXiv:hep-th/9412158],
A. Klemm, W. Lerche and S. Theisen, Int. J. Mod. Phys. A **11** (1996) 1929 [arXiv:hep-th/9505150].

[174] P. C. Argyres and A. E. Faraggi, Phys. Rev. Lett. **74** (1995) 3931 [arXiv:hep-th/9411057].

[175] P. C. Argyres and A. E. Faraggi, Phys. Rev. Lett. **74** (1995) 3931, [arXiv:hep-th/9411057].

[176] A. Hanany and Y. Oz, Nucl. Phys. B **452** (1995) 283, [arXiv:hep-th/9505075].

[177] P. C. Argyres, M. R. Plesser and A. D. Shapere, Phys. Rev. Lett. **75** (1995) 1699, [arXiv:hep-th/9505100].

[178] M. R. Douglas and S. H. Shenker, Nucl. Phys. B **447** (1995) 271 [arXiv:hep-th/9503163].

[179] D. Kutasov and A. Schwimmer, Phys. Lett. B **354** (1995) 315 [arXiv:hep-th/9505004].

[180] P. C. Argyres and M. R. Douglas, Nucl. Phys. B **448** (1995) 93 [arXiv:hep-th/9505062].

[181] A. Hanany and Y. Oz, Nucl. Phys. B **452**, 283 (1995) [arXiv:hep-th/9505075].

[182] P. C. Argyres, M. R. Plesser and A. D. Shapere, Phys. Rev. Lett. **75**, 1699 (1995) [arXiv:hep-th/9505100],
P. C. Argyres, A. E. Faraggi and A. D. Shapere, [arXiv:hep-th/9505190].

[183] J. A. Minahan and D. Nemeschansky, Nucl. Phys. B **464** (1996) 3 [arXiv:hep-th/9507032].

[184] O. Aharony and S. Yankielowicz, Nucl. Phys. B **473** (1996) 93 [arXiv:hep-th/9601011].

[185] J. A. Minahan and D. Nemeschansky, Nucl. Phys. B **468** (1996) 72 [arXiv:hep-th/9601059].

[186] P. C. Argyres, M. Ronen Plesser and N. Seiberg, Nucl. Phys. B **471** (1996) 159 [arXiv:hep-th/9603042].

[187] P. C. Argyres, M. Ronen Plesser, N. Seiberg and E. Witten, Nucl. Phys. B **461** (1996) 71 [arXiv:hep-th/9511154].

[188] P. C. Argyres and S. Pelland, JHEP **0003** (2000) 014 [arXiv:hep-th/9911255].

[189] N. B. Pomeroy, Phys. Lett. B **510** (2001) 305 [arXiv:hep-th/0103181].

[190] L. Alvarez-Gaume and S. F. Hassan, Fortsch. Phys. **45** (1997) 159 [arXiv:hep-th/9701069].

[191] S. V. Ketov, Fortsch. Phys. **45** (1997) 237 [arXiv:hep-th/9611209],
S. V. Ketov, [arXiv:hep-th/9710085].

[192] A. Bilal, [arXiv:hep-th/9601007],
A. Bilal, [arXiv:hep-th/0106246].

[193] W. Lerche, Nucl. Phys. Proc. Suppl. **55B** (1997) 83 [Fortsch. Phys. **45** (1997) 293] [arXiv:hep-th/9611190].

[194] R. Flume, L. O'Raifeartaigh and I. Sachs, [arXiv:hep-th/9611118].

[195] T. Eguchi, "Seiberg-Witten Theory And S-Duality," *talk given at NATO Advanced Study Institute on Strings, Branes and Dualities, Cargese, France, 26 May - 14 June, 1997.*

[196] Y. Ohta, [arXiv:hep-th/9903182].

[197] J. Thomassen, "Duality In N = 2 Supersymmetric Yang–Mills Theories," *Cand. Scient. Thesis, Niels Bohr Institute, University of Copenhagen, 1995.*

[198] C. Gomez and R. Hernandez, [arXiv:hep-th/9510023].

[199] E. Kiritsis, [arXiv:hep-ph/9911525].

[200] M. E. Peskin, [arXiv:hep-th/9702094].

[201] P. Di Vecchia, [arXiv:hep-th/9803026],
P. Di Vecchia, Surveys High Energ. Phys. **10** (1997) 119 [arXiv:hep-th/9608090].

[202] V. A. Novikov, M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **223** (1983) 445.

[203] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, M. B. Voloshin and V. I. Zakharov, Nucl. Phys. B **229** (1983) 394.

[204] D. Amati, K. Konishi, Y. Meurice, G. C. Rossi and G. Veneziano, Phys. Rept. **162** (1988) 169.

[205] D. Amati, G. C. Rossi and G. Veneziano, Nucl. Phys. B **249** (1985) 1, D. Amati, Y. Meurice, G. C. Rossi and G. Veneziano, Nucl. Phys. B **263**, 591 (1986), D. Amati and K. c. Chou, Phys. Lett. B **114** (1982) 129.

[206] J. Fuchs, Nucl. Phys. B **282**, 437 (1987).

[207] J. Fuchs, Nucl. Phys. B **272**, 677 (1986).

[208] J. Fuchs and M. G. Schmidt, Z. Phys. C **30**, 161 (1986).

[209] V. A. Novikov, M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **260**, 157 (1985) [Yad. Fiz. **42**, 1499 (1985)].

[210] V. A. Novikov, Sov. J. Nucl. Phys. **46**, 366 (1987) [Yad. Fiz. **46**, 656 (1987)].

[211] V. A. Novikov, Sov. J. Nucl. Phys. **46**, 554 (1987) [Yad. Fiz. **46**, 967 (1987)].

[212] D. Finnell and P. Pouliot, Nucl. Phys. B **453** (1995) 225 [arXiv:hep-th/9503115].

[213] N. Dorey, V. V. Khoze and M. P. Mattis, Phys. Rev. D **54** (1996) 2921 [arXiv:hep-th/9603136].

[214] N. Dorey, V. V. Khoze and M. P. Mattis, Phys. Lett. B **388** (1996) 324 [arXiv:hep-th/9607066].

[215] H. Aoyama, T. Harano, M. Sato and S. Wada, Phys. Lett. B **388** (1996) 331 [arXiv:hep-th/9607076].

[216] T. Harano and M. Sato, Nucl. Phys. B **484** (1997) 167 [arXiv:hep-th/9608060].

[217] N. Dorey, V. V. Khoze and M. P. Mattis, Phys. Rev. D **54** (1996) 7832 [arXiv:hep-th/9607202].

[218] K. Ito and N. Sasakura, Mod. Phys. Lett. A **12** (1997) 205 [arXiv:hep-th/9609104].

[219] N. Dorey, V. V. Khoze and M. P. Mattis, Nucl. Phys. B **492** (1997) 607 [arXiv:hep-th/9611016].

[220] N. Dorey, V. V. Khoze and M. P. Mattis, Phys. Lett. B **396** (1997) 141 [arXiv:hep-th/9612231].

[221] M. J. Slater, Phys. Lett. B **403** (1997) 57 [arXiv:hep-th/9701170].

[222] N. Dorey, T. J. Hollowood, V. V. Khoze and M. P. Mattis, Nucl. Phys. B **519** (1998) 470 [arXiv:hep-th/9709072],
N. Dorey, V. V. Khoze and M. P. Mattis, Nucl. Phys. B **513** (1998) 681 [arXiv:hep-th/9708036].

[223] V. V. Khoze, M. P. Mattis and M. J. Slater, Nucl. Phys. B **536** (1998) 69, [arXiv:hep-th/9804009].

[224] N. Dorey, T. J. Hollowood, V. V. Khoze and M. P. Mattis, [arXiv:hep-th/0206063].

[225] N. Dorey, T. J. Hollowood, V. V. Khoze, M. P. Mattis and S. Vandoren, Nucl. Phys. B **552** (1999) 88 [arXiv:hep-th/9901128],
N. Dorey, T. J. Hollowood, V. V. Khoze, M. P. Mattis and S. Vandoren, JHEP **9906** (1999) 023 [arXiv:hep-th/9810243],
N. Dorey, V. V. Khoze, M. P. Mattis and S. Vandoren, Phys. Lett. B **442** (1998) 145 [arXiv:hep-th/9808157].

[226] N. Dorey, V. V. Khoze, M. P. Mattis, M. J. Slater and W. A. Weir, Phys. Lett. B **408** (1997) 213 [arXiv:hep-th/9706007].

[227] M. J. Slater, Phys. Lett. B **481** (2000) 413 [arXiv:hep-th/9912011].

[228] T. J. Hollowood, V. V. Khoze and M. P. Mattis, Adv. Theor. Math. Phys. **4** (2000) 545 [arXiv:hep-th/9910118],
T. J. Hollowood, V. V. Khoze and M. P. Mattis, JHEP **9910** (1999) 019 [arXiv:hep-th/9905209].

[229] M. Matone, Phys. Lett. B **357** (1995) 342 [arXiv:hep-th/9506102].

[230] F. Fucito and G. Travaglini, Phys. Rev. D **55** (1997) 1099 [arXiv:hep-th/9605215].

[231] N. Dorey, V. V. Khoze and M. P. Mattis, Phys. Lett. B **390** (1997) 205 [arXiv:hep-th/9606199].

[232] T. Eguchi and S. K. Yang, Mod. Phys. Lett. A **11** (1996) 131 [arXiv:hep-th/9510183].

[233] E. D'Hoker, I. M. Krichever and D. H. Phong, Nucl. Phys. B **489** (1997) 179 [arXiv:hep-th/9609041].

[234] G. Chan and E. D'Hoker, Nucl. Phys. B **564** (2000) 503 [arXiv:hep-th/9906193],
G. J. Chan, [arXiv:hep-th/0106045],
G. Chan, [arXiv:hep-th/0105302].

[235] U. Lindstrom and M. Rocek, Phys. Lett. B **355** (1995) 492 [arXiv:hep-th/9503012].

[236] G. Bonelli, M. Matone and M. Tonin, Phys. Rev. D **55** (1997) 6466 [arXiv:hep-th/9610026].

[237] R. Flume, [arXiv:hep-th/9702192].

[238] R. Flume, M. Magro, L. O'Raifeartaigh, I. Sachs and O. Schnetz, Nucl. Phys. B **494** (1997) 331 [arXiv:hep-th/9611123],
M. Magro, L. O'Raifeartaigh and I. Sachs, Nucl. Phys. B **508** (1997) 433 [arXiv:hep-th/9704027].

[239] W. Nahm, [arXiv:hep-th/9608121].
W. Nahm, Nucl. Phys. Proc. Suppl. **58** (1997) 91.

[240] I. Sachs and W. Weir, Phys. Lett. B **479** (2000) 347 [arXiv:hep-th/9910015].

[241] U. Bruzzo, F. Fucito, A. Tanzini and G. Travaglini, Nucl. Phys. B **611** (2001) 205 [arXiv:hep-th/0008225].

[242] M. Matone, Phys. Lett. **514** (2001) 161 [arXiv:hep-th/0105041],
M. Matone, JHEP **0104** (2001) 041 [arXiv:hep-th/0103246].

[243] D. Bellisai, F. Fucito, A. Tanzini and G. Travaglini, JHEP **0007** (2000) 017 [arXiv:hep-th/0003272],
D. Bellisai, F. Fucito, A. Tanzini and G. Travaglini, Phys. Lett. B **480** (2000) 365 [arXiv:hep-th/0002110],
D. Bellisai, F. Fucito, A. Tanzini and G. Travaglini, [arXiv:hep-th/9812145].

[244] N. Dorey, V. V. Khoze and M. P. Mattis, Nucl. Phys. B **502** (1997) 94 [arXiv:hep-th/9704197],
N. Dorey, V. V. Khoze, M. P. Mattis, D. Tong and S. Vandoren, Nucl. Phys. B **502** (1997) 59 [arXiv:hep-th/9703228].

[245] N. Dorey, T. J. Hollowood and V. V. Khoze, [arXiv:hep-th/0010015].

[246] N. Dorey, T. J. Hollowood and V. V. Khoze, [arXiv:hep-th/0105090].

[247] F. Fucito, [arXiv:hep-th/0008220].

[248] F. Fucito, [arXiv:hep-th/9710160].

[249] H. M. Farkas and I. Kra, "Riemann Surfaces," *Graduate texts in mathematics; 71, New York: Springer, 2nd ed., 1992*.

[250] P. A. Griffiths, "Introduction to Algebraic Curves," *Translations of Mathematical Monographs Vol. 76, AMS, Providence, Rhode Island, 1989*.

[251] K. A. Intriligator and N. Seiberg, Nucl. Phys. B **444** (1995) 125 [arXiv:hep-th/9503179].

[252] P. C. Argyres, M. Ronen Plesser and N. Seiberg, Nucl. Phys. B **471** (1996) 159 [arXiv:hep-th/9603042].

[253] D. Kutasov, A. Schwimmer and N. Seiberg, Nucl. Phys. B **459** (1996) 455 [arXiv:hep-th/9510222].

[254] K. A. Intriligator and N. Seiberg, Nucl. Phys. Proc. Suppl. **45BC** (1996) 1 [arXiv:hep-th/9509066].

[255] N. Seiberg, Int. J. Mod. Phys. A **16** (2001) 4365 [arXiv:hep-th/9506077],
N. Seiberg, Prog. Theor. Phys. Suppl. **123** (1996) 337,
Int. J. Mod. Phys. A **12** (1997) 5171.

[256] N. Seiberg, Nucl. Phys. B **435** (1995) 129 [arXiv:hep-th/9411149].

[257] M. Hansen, "Duality and N = 1 Supersymmetric Gauge theories," *Cand. Scient. Thesis, Niels Bohr Institute, University of Copenhagen, 1998.*

[258] U. H. Danielsson and B. Sundborg, Phys. Lett. B **358** (1995) 273 [arXiv:hep-th/9504102].

[259] P. C. Argyres and A. D. Shapere, Nucl. Phys. B **461** (1996) 437 [arXiv:hep-th/9509175].

[260] P. C. Argyres, M. Ronen Plesser and A. D. Shapere, Nucl. Phys. B **483** (1997) 172 [arXiv:hep-th/9608129].

[261] A. Hanany, Nucl. Phys. B **466** (1996) 85 [arXiv:hep-th/9509176].

[262] U. H. Danielsson and B. Sundborg, Phys. Lett. B **370** (1996) 83 [arXiv:hep-th/9511180],
U. H. Danielsson and P. Stjernberg, Phys. Lett. B **380** (1996) 68 [arXiv:hep-th/9603082].

[263] J. A. Minahan, D. Nemeschansky and N. P. Warner, Nucl. Phys. B **528** (1998) 109 [arXiv:hep-th/9710146].

[264] J. A. Minahan, Nucl. Phys. B **537** (1999) 243 [arXiv:hep-th/9806246].

[265] A. Brandhuber and K. Landsteiner, Phys. Lett. B **358** (1995) 73 [arXiv:hep-th/9507008].

[266] M. Alishahiha, F. Ardalan and F. Mansouri, Phys. Lett. B **381** (1996) 446 [arXiv:hep-th/9512005].

[267] M. R. Abolhasani, M. Alishahiha and A. M. Ghezelbash, Nucl. Phys. B **480** (1996) 279 [arXiv:hep-th/9606043].

[268] K. Landsteiner, J. M. Pierre and S. B. Giddings, Phys. Rev. D **55** (1997) 2367 [arXiv:hep-th/9609059].

[269] T. Kitao, S. Terashima and S. K. Yang, Phys. Lett. B **399** (1997) 75 [arXiv:hep-th/9701009].

[270] P. C. Argyres, Adv. Theor. Math. Phys. **2** (1998) 293 [arXiv:hep-th/9706095],
P. C. Argyres and A. Buchel, Phys. Lett. B **431** (1998) 317 [arXiv:hep-th/9804007],

P. C. Argyres and A. Buchel, Phys. Lett. B **442** (1998) 180 [arXiv:hep-th/9806234], P. C. Argyres and A. Buchel, JHEP **9911** (1999) 014 [arXiv:hep-th/9910125].

[271] S. Terashima and S. K. Yang, Phys. Lett. B **430** (1998) 102 [arXiv:hep-th/9803014], S. Terashima and S. K. Yang, Nucl. Phys. B **537** (1999) 344 [arXiv:hep-th/9808022], J. Hashiba and S. Terashima, JHEP **9909** (1999) 020 [arXiv:hep-th/9909032].

[272] I. P. Ennes, S. G. Naculich, H. Rhedin and H. J. Schnitzer, Phys. Lett. B **452** (1999) 260 [arXiv:hep-th/9901124].

[273] I. P. Ennes, S. G. Naculich, H. Rhedin and H. J. Schnitzer, Nucl. Phys. B **558** (1999) 41 [arXiv:hep-th/9904078].

[274] E. Witten, Nucl. Phys. B **500** (1997) 3 [arXiv:hep-th/9703166].

[275] A. Giveon and D. Kutasov, Rev. Mod. Phys. **71** (4) (1999) 983 [arXiv:hep-th/9802067].

[276] N. D. Lambert and P. C. West, [arXiv:hep-th/9811177].

[277] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri and Y. Oz, Phys. Rept. **323** (2000) 183 [arXiv:hep-th/9905111].

[278] M. E. Peskin and D. V. Schroeder, "An Introduction To Quantum Field Theory," *Reading, U.S.A.: Addison-Wesley (1995).*

[279] S. Weinberg, "The Quantum Theory Of Fields. Vol. 3: Supersymmetry," *Cambridge, U.K.: Cambridge University Press (2000)* ,
S. Weinberg, "The Quantum Theory Of Fields. Vol. 2: Modern Applications," *Cambridge, U.K.: Cambridge University Press (1996)* ,
S. Weinberg, "The Quantum Theory Of Fields. Vol. 1: Foundations," *Cambridge, U.K.: Cambridge University Press (1995)* .

[280] G. 't Hooft, "Under The Spell Of The Gauge Principle," Adv. Ser. Math. Phys. **19** (1994) 1.

[281] E. S. Abers and B. W. Lee, Phys. Rept. **9** (1973) 1.

[282] T. Eguchi, P. B. Gilkey and A. J. Hanson, Phys. Rept. **66** (1980) 213.

[283] C. Nash, "Differential Topology And Quantum Field Theory," *London, U.K.: Academic Press, 1991.*

[284] G. M. Dixon, "Division Algebras: Octonions, Quaternions, Complex Numbers, And The Algebraic Design Of Physics," *Boston, U.S.A.: Kluwer Academic Publishers, 1994.*

[285] M. H. Freedman and F. Quinn, "Topology Of 4-manifolds," *Princeton, U.S.A.: Princeton University Press, 1990. Princeton Mathematical Series; 39*

[286] S. K. Donaldson, Topology **29** (1990) 257.

[287] S. K. Donaldson, Commun. Math. Phys. **96** (1984) 387.

[288] D. S. Freed and K. K. Uhlenbeck, "Instantons And Four-manifolds," *New York, U.S.A.: Springer-Verlag, 1984; Mathematical Sciences Research Institute publications; v. 1*

[289] H. B. Lawson and M. -L. Michelsohn, "Spin Geometry," *Princeton, U.S.A.: Princeton University Press, 1989; Princeton Mathematical Series; 38*

[290] S. K. Donaldson and P. B. Kronheimer, "The Geometry Of Four-manifolds," *Oxford, U.K.: Clarendon Press, 1990.*

[291] E. Witten, Math. Res. Lett. **1** (1994) 769 [arXiv:hep-th/9411102].

[292] M. Marino, [arXiv:hep-th/9701128].

[293] A. Klemm, [arXiv:hep-th/9705131].

[294] C. Nash, [arXiv:hep-th/9709135].

[295] R. Friedman, J. W. Morgan, eds., "Gauge Theory And The Topology Of Four-manifolds," *Providence, R.I., U.S.A.: American Mathematical Society, 1998; IAS/Park City Mathematics series; v. 4*

[296] G. Svetlichny, [arXiv:math-ph/9902027].

[297] K. Iga, [arXiv:hep-th/0207271].

[298] L. I. Nicolaescu, "Notes On Seiberg–Witten Theory," *Providence, U.S.A.: American Mathematical Society (2000)*

[299] M. Marcolli, [arXiv:dg-ga/9509005].